



Cloning Scheme for Generating pCMV-NS35

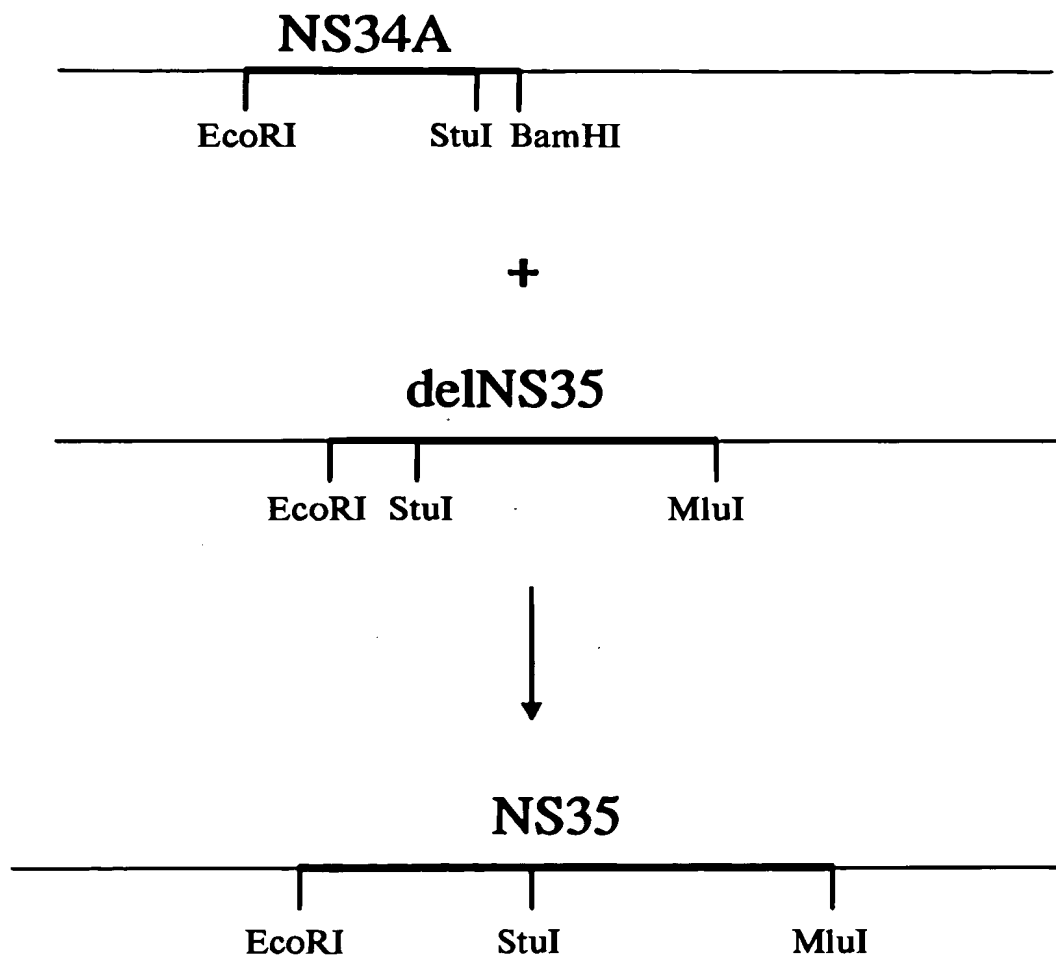


FIG. 1

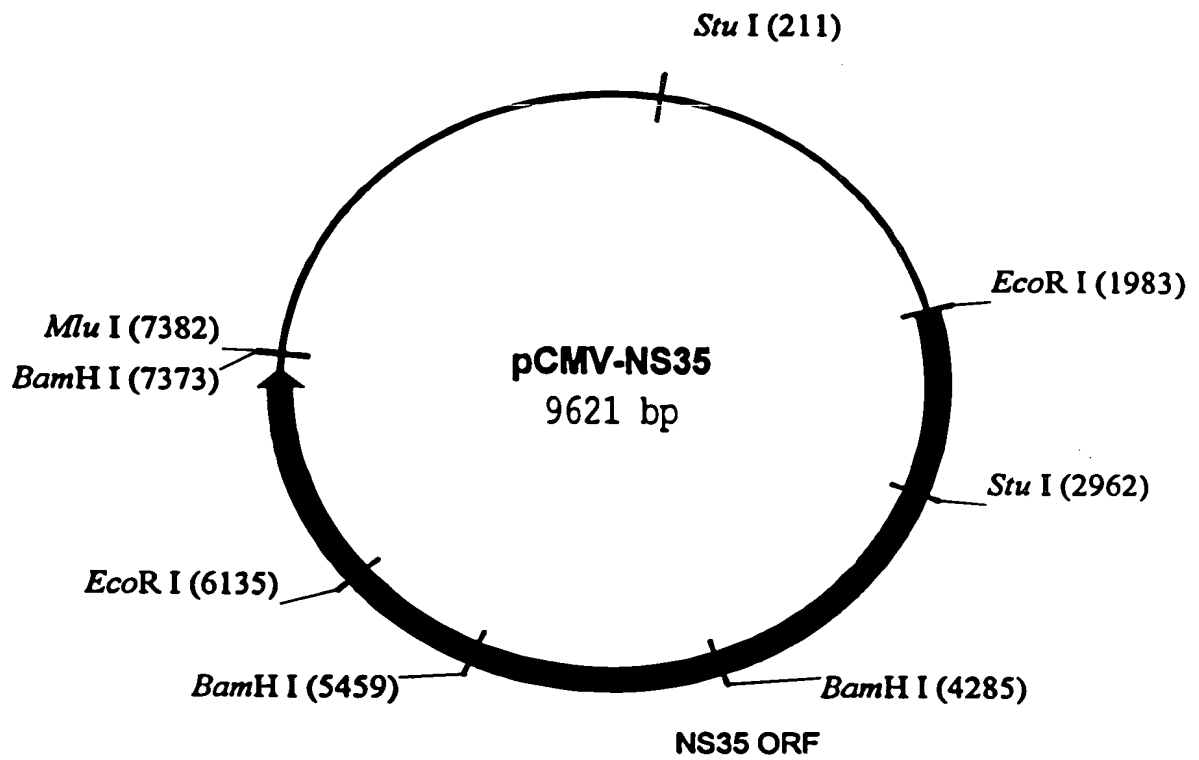


FIG. 2



pcMV-NS35

1 TC GCGCGGTTT CCGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG GAGACGGTCA CAGCTTGTCT GTAAGCGGAT
AGCGGCAAA GCCACTACTG CCACTTTTGG AGACTGTGTA CGTCGAGGCG CTCTGCCAGT GTCGAACAGA CATTCCGCTA

81 GCCGGGAGCA GACAAGCCCG TCAGCGGGTG TTGGCGGGTG TCGGGGCTGG CTTAACATATG CGGCATCAGA
CGGCCCTCGT CTGTTCCGGC AGTCCCGCGC AGTCGCCAC AACCGCCAC AGCCCCGACC GAATTGATAG GCGGTAGTCT

161 GCAGATTGTA CTGAGAGTGC ACCATATGAA GCTTTTTCGA AAAGCTAGG CCTCCAAAA AGCCTCCTCA CTACTTCTGG
CGTCTAACAT GACTCTCAG TGGTATACTT CGAAAAACGT TTTCCGATCC GGAGGTTTT TCGGAGGAGT GATGAAGACC

241 AATAGCTCAG AGGCCGAGGC GGCCTCGGCC TCTGCATAAA TAAAAAAAT TAGTCAGCCA TGGGCGGGAG AATGGCGGA
TTATCGAGTC TCCGGCTCCG CCGGAGCCGG AGAGTATTT ATTTTTTTTA ATCAGTCGGT ACCCGCCTC TTACCCGCCT

321 ACTGGGCGG GAGGGAATTA TTGGCTATTG GCCATTGCAT ACGTTGTATC TATATCATAA TATGTACATT TATATTGGCT
TGACCGGCCC CTCCCTTAAT AACCGATAAC CGGTAACGTA TGCAACATAG ATATAGTATT ATACATGTAA ATATAACCGA

401 CATGTCCAAT ATGACCGCCA TGTGACATT GATTATTGAC TAGTTATTAA TAGTAATCAA TTACGGGGTC ATTAGTTTCA
GTACAGGTTA TACTGGCGGT ACAACTGTAA CTAATAACTG ATCAATAATT ATCATTAGTT AATGCCCCAG TAATCAAGTA

481 AGCCCATATA TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC CCGCCCCATT
TCGGGTATAT ACCTCAAGGC GCAATGTATT GAATGCCATT TACCGGGCGG ACCGACTGC GGGTTGCTGG GGGCGGGTAA

561 GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA GGCACTTTCC ATTGACGTCA ATGGGTGGAG TATTACGGT
CTGCAGTTAT TACTGCATAC AAGGGTATCA TTGCGGTTAT CCCTGAAAGG TAACTGCAGT TACCCACCTC ATAAATGCCA

641 AAAGTCCCCA CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTCGGCCC CCTATTGACG TCAATGACGG TAAATGGCCC
TTTGACGGGT GAACCGTCAT GTAGTTACA TAGTATACGG TTCAGGGCGG GGATAACTGC AGTACTGCC ATTTACCGGG

StuI

FIG. 3A



pCMV-NS35

721 GCCTGGCATT ATGCCCAGTA CATGACCTTA CGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA TCGCTATTAC
CGGACCGTAA TACGGGTCAAT GTACTGGAAT GCCCTGAAAG GATGAACCGT CATGTAGATG CATAATCAGT AGCGATAATG

801 CATGGTGATG CCGTTTTTGGC AGTACACCAA TGGGCGTGGA TAGCGGTTTG ACTCACGGGG ATTTCCAAGT CTCCACCCCA
GTACCACTAC GCCAAAACCG TCATGTGGTT ACCCGCACCT ATCGCCAAAC TGAGTGCCCC TAAAGGTTCA GAGGTGGGGT

881 TTGACGTCAA TGGGAGTTTG TTTTGGCACC AAAATCAACG GGACTTTCCA AAATGTCGTA ATAAACCCCGC CCCGTTGACG
AACTGCAGTT ACCCTCAAAC AAAACCGTGG TTTTAGTTGC CCTGAAAGGT TTTACAGCAT TATTGGGGCG GGGCAACTGC

961 CAAATGGCG GTAGCGGTGT ACGTGGGAG GTCTATATAA GCAGAGCTCG TTTAGTGAAC CGTCAGATCG CCTGGAGACG
GTTTACCCGC CATCCGCACA TGCCACCCCTC CAGATATATT CGTCTCGAGC AAATCACTTG GCAGTCTAGC GGACCTCTGC

1041 CCATCCACGC TGTTTTGACC TCCATAGAAG ACACGGGAC CGATCCAGCC TCCGCGGGCG GGAACGGTGC ATTGGAACGC
GGTAGGTGG ACAAACCTGG AGGTATCTTC TGTGGCCCTG CCTAGGTCCG AGCGCCCGG CCTTGCCACG TAACCTTGCG

1121 GGATTCCCG TGCCAAGAGT GACGTAAGTA CCGCCTATAG ACTCTATAGG CACACCCCTT TGGCTCTTAT GCATGCTATA
CCTAAGGGGC ACGGTTCTCA CTGCATTTCAT GCGGATATC TGAGATATCC GTGTGGGAA ACCGAGAATA CGTACGATAT

1201 CTGTTTTTGG CTTGGGGCCT ATACACCCCC GCTCCTTATG CTATAGGTGA TGGTATAGCT TAGCCTATAG GTGTGGGTTA
GACAAAACC GAACCCCGGA TATGTGGGG CGAGGAATAC GATATCCACT ACCATATCGA ATCGGATATC CACACCCAAT

1281 TTGACCATTA TTGACCACTC CCCTATTGGT GACGATACTT TCCATTACTA ATCCATAACA TGGCTCTTTG CCACAACCTAT
AACTGGTAAT AACTGGTGAG GGGATAACCA CTGCTATGAA AGGTAATGAT TAGGTATTGT ACCGAGAAAC GGTGTTGATA

1361 CTCTATTGGC TATATGCCAA TACTCTGTCC TTCAGAGACT GACACGGACT CTGTATTTTT ACAGGATGGG GTCCATTAT
GAGATAACCG ATATACGGTT ATGAGACAGG AAGTCTCTGA CTGTGCCTGA GACATAAAAA TGTCTTACCC CAGGTAAATA

FIG. 3B



pCMV-NS35

```
1441 TATTACAAA TTCACATATA CAACAACGCC GTCCCCCGTG CCCGCAGTTT TTATTAAACA TAGCGTGGGA TCTCCGACAT
    ATAAATGTTT AAGTGATAT GTTGTGCGG CAGGGGGCAC GGGCGTCAAA AATAATTGT ATCGCACCT AGAGGCTGTA

1521 CTCGGGTACG TGTTCCGGAC ATGGGCTCTT CTCCGGTAGC GGGGGAGCTT CCACATCCGA GCCCTGGTCC CATCCGTCCA
    GAGCCCATGC ACAAGGCCTG TACCCGAGAA GAGGCCATCG CCGCCTCGAA GGTGTAGGCT CGGGACCAGG GTAGGCAGGT

1601 GCGGCTCATG GTCGCTCGG AGCTCCTTGC TCCTAACAGT GGAGGCCAGA CTTAGGCACA GCACAATGCC CACCACCACC
    CGCCGAGTAC CAGCGAGCCG TCGAGGAACG AGGATTGTCA CCTCCGGTCT GAATCCGTGT CGTGTACGG GTGGTGGTGG

1681 AGTGTGCGC ACAAGGCCGT GCGGCTAGG TATGTGTCTG AAAATGAGCT CGGAGATTGG GCTGGCACCT GGACGCAGAT
    TCACACGGCG TGTTCGGCA CCGCCATCCC ATACACAGAC TTTTACTCGA GCCTCTAACC CGAGCGTGGA CTTGCGTCTA

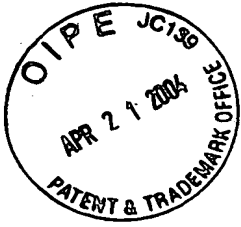
1761 GGAAGACTTA AGCAGCGGC AGAAGAAGAT GCAGGCAGCT GAGTTGTTGT ATTCTGATAA GAGTCAGAGG TAACTCCCGT
    CCTTCTGAAT TCCGTGCGC TCTTCTTCTA CGTCCGTGCA CTCAACAACA TAAGACTATT CTCAGTCTCC ATTGAGGGCA

1841 TCGGCTGCTG TTAACGGTGG AGGCAGTGT AGTCTGAGCA GTACTCGTTG CTGCCGGCGG CGCCACCAGA CATAATAGCT
    AGCCACGGAC AATTGCCACC TCCCGTCACA TCAGACTCGT CATGAGCAAC GACGGCGCGC GCGGTGGTCT GTATTATCGA

+2 M A A
    EcoRI
    -----
1921 GACAGACTAA CAGACTGTTT CTTTCCATGG GTCTTTTCTG CAGTCACCGT CGTCGACCTA AGAATTCACC ATGGCTGCAT
    CTGTCTGATT GTCTGACAAG GAAAGGTACC CAGAAAAGAC GTCAGTGCCA GCAGCTGGAT TCTTAAGTGG TACCGACGTA

+2 Y A A Q G Y K V L V L N P S V A A T L G F G A Y M S K
2001 ATGCAGCTCA GGGCTATAAG GTGCTAGTAC TCAACCCCTC TGTTGCTGCA ACACTGGGT TTGGTGCTTA CATGTCCAAG
    TAGCTCGAGT CCCGATATTC CACGATCATG AGTTGGGGAG ACAACGACGT TGTGACCCGA AACCCAGAAAT GTACAGGTTT
```

FIG. 3C



pCMV-NS35

+2 A H G I D P N I R T G V R T I T T G S P I T Y S T Y G
2081 GCTCATGGGA TCGATCCTAA CATCAGGACC GGGGTGAGAA CAATTACCAC TGGCAGCCCC ATCAGTACT CCACCTACGG
CGAGTACCCT AGCTAGGATT GTAGTCTTGG CCCCACTCTT GTTAATGGTG ACCGTGGGG TAGTGATGA GGTGGATGCC

+2 K F L A D G G C S G G A Y D I I I C D E C H S T D A
2161 CAAGTTCCTT GCCGACGGCG GGTGCTCGG GGGCGCTTAT GACATAATAA TTTGTGACGA GTGCCACTCC ACGGATGCCA
GTTCAGGAA CGGCTGCCGC CCACGAGCCC CCCGGGAATA CTGTATTATT AAACACTGCT CACGGTGAGG TGCTACGGT

+2 T S I L G I G T V L D Q A E T A G A R L V V L A T A T
2241 CATCCATCTT GGGCATTGGC ACTGCTCTG ACCAAGCAGA GACTGCGGG GCGAGACTGG TTGTGCTCGC CACCGCCACC
GTAGGTAGAA CCCGTAACCG TGACAGGAAC TGGTTCGTCT CTGACGCCCC CGCTCTGACC AACACGAGCG GTGGCGGTGG

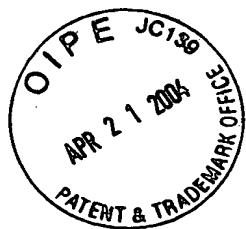
+2 P P G S V T V P H P N I E E V A L S T T G E I P F Y G
2321 CCTCCGGCT CCGTCACTGT GCGCCATCCC AACATCGAG AGTTGCTCT GTCCACCACC GGAGAGATCC CTTTTACGG
GGAGGCCCCA GGCAGTGACA CCGGTAGGG TTGTAGCTCC TCCAACGAGA CAGGTGGTGG CCTCTCTAGG GAAAAATGCC

+2 K A I P L E V I K G G R H L I F C H S K K K C D E L
2401 CAAGGCTATC CCCCTCGAAG TAATCAAGGG GGGGAGACAT CTCATCTTCT GTCATTCAAA GAAGAAGTGC GACGAACCTCG
GTTCCGATAG GGGGAGCTTC ATTAGTTCCC CCCCTCTGTA GAGTAGAAGA CAGTAAGTTT CTCTCTCAGG CTGCTTGAGC

+2 A A K L V A L G I N A V A Y Y R G L D V S V I P T S G
2481 CCGCAAAGCT GGTGCGATTG GGCATCAATG CCGTGGCCTA CTACCGCGGT CTTGACGTGT CCGTCATCCC GACCAGCGGC
GGCGTTTCCA CCAGCGTAAC CCGTAGTTAC GGCACCGGAT GATGGCGCCA GAACTGCACA GGCAGTAGGG CTGGTCGCCC

+2 D V V V A T D A L M T G Y T G D F D S V I D C N T C
2561 GATGTTGTCG TCGTGGCAAC CGATGCCCTC ATGACCGGCT ATACCGCGGA CTTCCGACTCG GTGATAGACT GCAATACGTG
CTACAACAGC AGCACCGGTG GCTACGGGAG TACTGGCCGA TATGGCCGCT GAAGCTGAGC CACTATCTGA CGTTATGCAC

FIG. 3D



pCMV-NS35

```
+2 V T Q T V D F S L D P T F T I E T I T L P Q D A V S
2641 TGTACCCAG ACAGTCGATT TCAGCCTTGA CCCTACCTTC ACCATTGAGA CAATCAGGCT CCCCCAAGAT GCTGTCTCCC
ACAGTGGGTC TGTACGCTAA AGTCGGAAC TGGTGAAG TGGTAACTCT GTTAGTGCGA GGGGGTTCTA CGACAGAGGG

+2 R T Q R R G R T G R G K P G I Y R F V A P G E R P S G
2721 GCACTCAACG TCGGGGCAGG ACTGGCAGG GGAAGCCAGG CATCTACAGA TTGTGGCAC CGGGGGAGCG CCCCTCCGGC
CGTGAGTTGC AGCCCCGTCC TGACCGTCCC CCTTCGGTCC GTAGATGCT AAACACCGTG CCCCCCTCGC GGGGAGGCGG

+2 M F D S S V L C E C Y D A G C A W Y E L T P A E T T V
2801 ATGTTGACT CGTCCGTCTT CTGTGAGTGC TATGACGCAG GCTGTGCTTG GTATGAGCTC ACGCCCGCCG AGACTACAGT
TACAAGCTGA GCAGGCAGGA GAACTCAGG ATACTGCGTC CGACACGAAC CATACTCGAG TCGGGGCGG TCTGATGTCA

+2 R L R A Y M N T P G L P V C Q D H L E F W E G V F T
StuI

2881 TAGGCTACGA GCGTACATGA ACACCCCGGG GCTTCCCGTG TGCCAGGACC ATCTTGAATT TTGGGAGGGC GTCTTTACAG
ATCCGATGCT CGCATGTACT TGTGGGGCCC CGAAGGGCAC ACGTCTCTGG TAGAACTTAA AACCTTCCCG CAGAAATGTC

+2 G L T H I D A H F L S Q T K Q S G E N L P Y L V A Y Q
StuI
----
2961 GCCTCACTCA TATAGATGCC CACTTTCTAT CCCAGACAAA GCAGAGTGGG GAGAACCTTC CTTACCTGGT AGCGTACCAA
CGGAGTGAGT ATATCTACGG GTGAAAGATA GGGTCTGTTT CGTCTCAGCC CTCTTGAAG GAATGGACCA TCGCATGGTT

+2 A T V C A R A Q A P P P S W D Q M W K C L I R L K P T
3041 GCCACCGTGT GCGTAGGGC TCAAGCCCCCT CCCCCATCGT GGGACAGAT GTGGAAGTGT TTGATTGCGC TCAAGCCAC
CGGTGGCACA CGCGATCCCG AGTTCGGGA GGGGGTAGCA CCCTGGTCTA CACCTTCACA AACTAAGCGG AGTTCGGGTG
```

FIG. 3E



pCMV-NS35

+2 L H G P T P L L Y R L G A V Q N E I T L T H P V T K
3121 CCTCCATGG CCAACACCCC TGCTATACAG ACTGGGCGCT GTTCAGATG AATCACCCT GACGCACCCA GTCACCAAAAT
GGAGGTACCC GGTGTGGG ACATATGTC TGACCCGGA CAAGTCTTAC TTATGTGGA CTGGGTGGT CAGTGGTTA

+2 Y I M T C M S A D L E V V T S T W V L V G G V L A A L
3201 ACATCATGAC ATGCATGTC GCGACCTGG AGTCTGTAC GAGCACCTGG GTGCTCGTTG GCGGCTCCT GGTGCTTTG
TGTACTACTG TACGTACAGC CGGCTGGACC TCCAGCAGTG CTCGTGGACC CACGAGCAAC CGCCGCAGGA CGGACGAAAC

+2 A A Y C L S T G C V V I V G R V V L S G K P A I I P D
3281 GCGGCTATT GCTGTCAAC AGGCTGCGTG GTCATAGTGG GCAGGCTCGT CTTGTCCGG AAGCCGGCAA TCATACCTGA
CGGCGCATAA CGACAGTTG TCCGACGCAC CAGTATCACC CGTCCAGCA GAACAGGCC TTCCGCCGT AGTATGGAAT

+2 R E V L Y R E F D E M E E C S Q H L P Y I E Q G M M
3361 CAGGGAATC CTCIACCGAG AGTTCGATGA GATGAAGC TGCTTCAGC ACTTACCGTA CATCGAGCAA GGGATGATGC
GTCCCTTCAG GAGATGGCTC TCAAGCTACT CTACCTTCTC ACGAGAGTCG TGAATGGCAT GTAGCTCGTT CCCTACTACG

+2 L A E Q F K Q K A L G L L Q T A S R Q A E V I A P A V
3441 TCGCCGAGCA GTTCAAGCAG AAGGCGCTCG GCTCTGCA GACCGCTCC CGTCAGGCAG AGTTATCGC CCCTGCTGTC
AGCGGCTCGT CAAGTTCGTC TTCCGGGAGC CGGAGGACGT CTGGCGCAGG GCAGTCCGTC TCCAATAGCG GGGACGACAG

+2 Q T N W Q K L E T F W A K H M W N F I S G I Q Y L A G
3521 CAGACCAACT GCGAAAACT CGAGACCTTC TGGGCGAAGC ATATGTGAA CTTCAATCAGT GGGATACAAT ACTTGGCGGG
GTCTGTTGA CCGTTTTTGA GCTCTGGAAG ACCCGCTTCG TATACACCTT GAAGTAGTCA CCCTATGTTA TGAACCGCCC

+2 L S T L P G N P A I A S L M A F T A A V T S P L T T
3601 CTTGTCAACG CTGCTGTA ACCCGGCAT TCGTTCATG ATGGTTTGA CAGCTGCTGT CACCGAGCCA CTAACCACTA
GAACAGTTGC GACGGACCAT TGGGGGGTA ACGAAGTAAC TACCGAAAAAT CTCGACGACA GTGGTGGGT GATTGGTGAT

FIG. 3F



pCMV-NS35

+2 S Q T L L F N I L G G W V A A Q L A A P G A A T A F V
3681 GCCAAACCT CCTCTTCAAC ATATTGGGG GGTGGTGGC TGCCAGCTC GCCGCCCG GTGCGGTAC TGCCTTTGTG
CGGTTTGGG GGAGAAATTG TATAACCCC CCACCCACG ACGGTGAG CGCGGGGGC CACGGCGATG ACGAAACAC

+2 G A G L A G A A I G S V G L G K V L I D I L A G Y G A
3761 GCGCTGGCT TAGTGGCG CGCCATCGG AGTGTGGAC TGGGAAGGT CCTCATAGAC ATCCTTGCAG GGTATGGCG
CCGACCGA ATCGACCG GCGTAGCCG TCACAACCTG ACCCTTCCA GGAGTATCTG TAGGAACGTC CCATACCGG

+2 G V A G A L V A F K I M S G E V P S T E D L V N L L
3841 GGGCTGGC GGAGCTCTTG TGGCATTCAA GATCATGAG GGTGAGGTCC CCTCCACGGA GGACTGGTC AATCTACTGC
CCGACCGC CTTGAGAAC ACCGTAACTT CTAGTACTCG CCACTCCAGG GGAGTGCCT CCTGGACGAG TTAGATGAG

+2 P A I L S P G A L V V G V V C A A I L R R H V G P G E
3921 CCGCATCCT CTGCCCCGA GCCCTCGTAG TCGGCTGCT CTGTGAGCA ATACTGGCC GGCACGTTGG CCGGGCGGAG
GGCGTAGGA GAGCGGCCT CGGAGCATC AGCCGACCA GACAGTCTGT TATGACGCG CGTGCAACC GGGCCCGCTC

+2 G A V Q W M N R L I A F A S R G N H V S P T H Y V P E
4001 GGGCAGTGC AGTGGATGAA CCGCTGATA GCCTTCGCT CCGGGGGA CCATGTTTC CCCACGCACT ACGTGCCGGA
CCCCGTCAG TCACTACTT GCGCGACTAT CGAAGCGGA GGGCCCCCTT GGTACAAAG GGTGCGTGA TGCACGCT

+2 S D A A A R V T A I L S S L T V T Q L L R R L H Q W
4081 GAGCGATGCA GTCGCCCG TCACTGCCAT ACTCAGCAG CTCAGTGTA CCCAGTCTT GAGCGGACTG CACCACTGGA
CTCGCTAGT CGACGGCGC AGTGACGGTA TGAGTCTCG GAGTGACATT GGTGCGAGGA CTCGCTGAC GTGTCACCT

+2 I S S E C T T P C S G S W L R D I W D W I C E V L S D
4161 TAAGCTCGA GTGTACCACT CCATGCTCG GTTCTGGCT AAGGACATC TGGGACTGGA TATCGGAGT GTTGAGCGAC
ATTCGAGCCT CACATGGTGA GTTACGAGG CAAGGACCGA TTCCCTGTAG ACCCTGACCT ATACGCTCCA CAACCTCGCT

FIG. 3G



pCMV-NS35

+2 F K T W L K A K L M P Q L P G I P F V S C Q R G Y K G
BamHI

4241 TTTAAGACCT GGCTAAAAGC TAAGCTCATG CCACAGCTGC CTGGGATCCC CTTTGTGTCC TGCCAGCGCG GGTATAAGGG
AAATTCTGGA CCGATTTCG ATTGAGTAC GGTTCGACG GACCTAGGG GAAACACAGG ACGGTGCGC CCATATTCCC
+2 V W R G D G I M H T R C H C G A E I T G H V K N G T
4321 GGTCTGGCGA GGGACGGCA TCATGCACAC TCGTGGCAC TGTGGAGCTG AGATCACTGG ACATGTCAA AACGGGACGA
CCAGACCGCT CCCCTGCCGT AGTACGTGT AGCGACGGTG ACACCTGCAC TCTAGTGACC TGTACAGTTT TTGCCCTGCT
+2 M R I V G P R T C R N M W S G T F P I N A Y T T G P C
4401 TGAGGATCGT CCGTCTTAGG ACCTGCAGGA ACATGTGGAG TGGGACCTTC CCCATTAAATG CCTACACCAC GGGCCCTGT
ACTCCTAGCA GCCAGGATCC TGGACGTCTT TGTACACCTC ACCCTGGAAG GGTAATTAC GGATGTGGTG CCCGGGGACA
+2 T P L P A P N Y T F A L W R V S A E E Y V E I R Q V G
4481 ACCCCCTTC CTGCGCCGAA CTACACGTC GCGTATGGA GGTGTCTGC AGAGGAATAC GTGGAGATAA GGCAGGTGG
TGGGGGAAG GACGGGCTT GATGTGCAAG CCGATACCT CCCACAGACG TCTCCTTATG CACCTCTATT CCGTCCACCC
+2 D F H Y V T G M T T D N L K C P C Q V P S P E F T
4561 GGAATCCAC TACGTGACGG GTATGACTAC TGACAATCTT AAATGCCGT GCCAGTCCC ATCGCCCGAA TTTTTCACAG
CCTGAAGGTG ATGCACTGCC CATACTGATG ACTGTTAGAA TTTACGGGCA CCGTCCAGGG TAGCGGCTT AAAAAGTGC
+2 E L D G V R L H R F A P P C K P L L R E E V S F R V G
4641 AATTGGACGG GTGCGCCTA CATAGTTTG CCCCCCTTG CAAGCCCTTG CTGCGGAGG AGGTATCATT CAGAGTAGGA
TTAACCTGCC CCACGGGAT GTATCCAAAC GCGGGGGGAC GTTCGGGAC GACGCCCTCC TCCATAGTAA GTCTCATCT
+2 L H E Y P V G S Q L P C E P E P D V A V L T S M L T D
4721 CTCCACGAAT ACCCGGTAGG GTCGCAATTA CCTTGGGAG CCGAACCGGA CGTGGCGGTG TTGACGTCCA TGCTCACTGA
GAGGTGCTTA TGGGCCATCC CAGCGTTAAT GGAACGCTCG GGCTTGGCT GCACGGCAC AACTGCAGGT ACGAGTGACT

FIG. 3H



pCMV-NS35

+2 P S H I T A E A A G R R L A R G S P P S V A S S S A
4801 TCCCTCCCAT ATAACAGCAG AGCGGCCGG GCGAAGGTG CCGAGGGGAT CACCCCTC TGIGGCCAGC TCCTCGGCTA
AGGAGGGTA TATTGTCGTC TCCGCCGGCC CGCTTCCAAC CGCTCCCTA GTGGGGGAG ACACCGGTG AGAGCCGAT

+2 S Q L S A P S L K A T C T A N H D S P D A E L I E A N
4881 GCCAGCTATC CGCTCCATCT CTCAGGCAA CTTGCACCGC TAACCATGAC TCCCTGATG CTGAGCTCAT AGAGCCCAAC
CGGTGATAG GCGAGGTAGA GAGTCCGTT GAACGTGGC ATTGTTACTG AGGGACTAC GACTCGAGTA TCTCCGGTTG

+2 L L W R Q E M G G N I T R V E S E N K V V I L D S F D
4961 CTCCTATGGA GCGAGGAGT GCGCGGCAAC ATCACCAGG TTGAGTCAGA AACAAAGTG GTGATTTCTGG ACTCCTTGA
GAGGATACCT CCGTCTCTA CCGCGCGTTG TAGTGGTCCC AACTCAGTCT TTTGTTTAC CACTAAGACC TGAGGAAGCT

+2 P L V A E E D E R E I S V P A E I L R K S R R F A Q
5041 TCCGCTTGT GCGGAGGAG ACAGCGGA GATCTCCGA CCGCGAGAA TCCTGGGAA GTCTCGGAGA TTGCCCCAGG
AGCGAACAC CGCTCTCTC TGCTGCCCT CTAGAGGCAT GCGGTCTTT AGGACGCCCT CAGAGCCTCT AAGCGGGTCC

+2 A L P V W A R P D Y N P P L V E T W K K P D Y E P P V
5121 CCCTGCCCCG TTGGCGCGG CCGGACTATA ACCCCCGCT AGTGAGAGC TGGAAAAGC CCGACTACGA ACCACCTGTG
GGGACGGGCA AACCGGCGC GGCCTGATAT TGGGGGGCA TCACCTCTGC ACCTTTTTCG GGTGATGCT TGGTGGACAC

+2 V H G C P L P P P K S P P V P P P R K K R T V V L T E
5201 GTCCATGGT GCCCGTTC ACCTCAAAG TCCCTCTCTG TGCTCCGCC TCGGAAGAAG CCGAGGGTGG TCCTCACTGA
CAGGTACCGA CCGGCGAAG TGGAGTTTC AGGGAGGAC ACGGAGCGG AGCCTTCTTC GCCTGCCACC AGGAGTGACT

+2 S T L S T A L A E L A T R S F G S S S T S G I T G D
5281 ATCAACCCTA TCTACTGGCT TGGCCGAGCT CGCCACCAGA AGCTTTGGCA GCTCCTCAAC TTCGGGCATT ACGGGCGACA
TAGTTGGGAT AGATGACGA ACCGGCTCGA GCGGTGGTCT TCGAAACCGT CGAGGAGTTG AAGGCCGTAA TGCCCCGTGT

FIG. 3I



pCMV-NS35

+2 N T T S S E P A P S G C P P D S D A E S Y S S M P P
5361 ATACGACAAC ATCCTCTGAG CCGGCCCTT CTGGCTGCCC CCCGACTCC GACGCTGAGT CCTATTCTTC CATGCCCCCC
TATGCTGTTG TAGGAGACTC GGGCGGGGAA GACCGACGGG GGGGCTGAGG CTGCGACTCA GGATAAGGAG GTACGGGGGG

+2 L E G E P G D P D L S D G S W S T V S S E A N A E D V
BamHI

5441 CTGGAGGGG AGCCTGGGA TCCGGATCTT AGCGACGGGT CATGGTCAAC GGTCACTAGT GAGGCCAAG CCGAGGATGT
GACCTCCCC TCGGACCCCT AGGCCTAGAA TCGCTGCCCC GTACCAGTTG CCAGTCATCA CTCCGGTTGC GCCTCCTACA

+2 V C C S M S Y S W T G A L V T P C A A E E Q K L P I
5521 CGTGTGCTGC TCAATGTCTT ACTCTTGGAC AGCGGCACTC GTCACCCCGT GCGCGGGGA AGAACAGAAA CTGCCCATCA
GCACACGAG AGTTACAGAA TGAGAACCTG TCCCGGTGAG CAGTGGGGCA CCGGCGCCCT TCTTGTCTTT GACGGGTAGT

+2 N A L S N S L L R H H N L V Y S T T S R S A C Q R Q K
5601 ATGCACTAAG CAACTCGTTG CTACGTCACC ACAATTGGT GTATTCCACC ACCTCAGGCA GTGCTTGCCA AAGGCAGAAG
TACGTGATTC GTTGAGCAAC GATGCAGTGG TGTTAAACCA CATAAGGTGG TGGACTGGT CACGAACGGT TTCCGTCTTC

+2 K V T F D R L Q V L D S H Y Q D V L K E V K A A A S K
5681 AAAGTCACAT TTGACAGACT GCAAGTTCTG GACAGCCATT ACCAGGACGT ACTCAAGGAG GTTAAAGCAG CGGCGTCAAA
TTTCAGTCTA AACTGTCTGA CGTTCAAGAC CTGTGCTGTA TGGTCTGCA TGAGTCTCTC CAATTTCGTC GCCGCAGTTT

+2 V K A N L L S V E E A C S L T P P H S A K S K F G Y
5761 AGTGAAGGCT AACTTGCTAT CCGTAGAGGA AGCTTGACGC CTGACGCCCC CACACTCAGC CAAATCCAAG TTGTTTATG
TCACCTCCGA TTGAACGATA GGCATCTCCT TCGAACGTCG GACTGCGGGG GTGTGAGTCG GTTTAGGTTT AAACCAATAC

FIG. 3J



pCMV-NS35

+2 G A K D V R C H A R K A V T H I N S V W K D L L E D N
5841 GGGAAAAGA CGTCCGTTGC CATGCCAGAA AGCCGTAAC CCACATCAAC TCCGTGTGGA AAGACCTTCT GGAAGACAAAT
CCCCTTTTCT GCAGGCAACG GTACGCTCTT TCCGGCATTG GGTGTAGTGG AGGCACACCT TTCTGGAAGA CCTTCTGTTA

+2 V T P I D T T I M A K N E V F C V Q P E K G R K P A
5921 GTAACACCAA TAGACACTAC CATCATGGCT AAGAACGAGG TTTTCTGCGT TCAGCTGAG AAGGGGGTTC GTAAGCCAGC
CATTCTGGTT ATCTGTGATG GTAGTACCGA TTCTTGCTCC AAAAGACGCA AGTCGGACTC TTCCCCCCAG CATTCGGTCC

+2 R L I V F P D L G V R V C E K M A L Y D V V T K L P
6001 TCCTCTCATC GTGTTCCCG ATCTGGCGT GCGGTGTGC GAAAGATGG CTTTGTACGA CGTGGTTACA AAGCTCCCT
AGCAGAGTAG CACAAGGGGC TAGACCGCA CGCGCACACG CTTTTTCTACC GAAACATGCT GCACCAATGT TTCCGAGGGGA

+2 L A V M G S S Y G F Q Y S P G Q R V E F L V Q A W K S
EcoRI

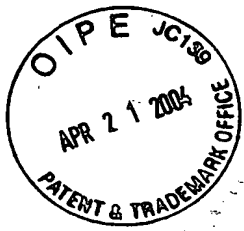
6081 TGGCCGTGAT GGAAGCTCC TAGGATTCC AATACTCACC AGGACAGCG GTTGAATTCC TCGTGCAAGC GTGGAAGTCC
ACCGGCACTA CCTTCGAGG ATGCCAAGG TTATGAGTGG TCCTGTCGCC CAACTTAAGG AGCAGCTTCG CACCTTCAGG

+2 K K T P M G F S Y D T R C F D S T V T E S D I R T E E
6161 AAGAAAACCC CAATGGGTT CTGCTATGAT ACCGCTGCT TTGACTCCAC AGTCACTGAG AGGCACATCC GTACGGAGGA
TTCTTTGGG GTTACCCCAA GAGCATACTA TGGCGACGA AACTGAGGTG TCAGTGACTC TCGCTGTAGG CATGCCCTCT

+2 A I Y Q C C D L D P Q A R V A I K S L T E R L Y V G
6241 GGCAATCTAC CAATGTTGTG ACCTCGACCC CCAAGCCCGC GTGGCCATCA AGTCCCTCAC CGAGAGGCTT TATGTTGGGG
CCGTTAGATG GTTACAACAC TGGAGCTGGG GGTTCGGCGG CACCGGTAGT TCAGGGAGTG GCTCTCCGAA ATACAACCCC

+2 G P L T N S R G E N C G Y R R C R A S G V L T T S C G
6321 GCCCTCTTAC CAATTCAAGG GGGGAGAACT GCGGTATCG CAGGTGCCG CCGAGCGCG TACTGACAAC TAGCTGTGGT
CGGAGAAATG GTTAAGTTCC CCCCTTTGA CGCCGATAGC GTCCACGGCG CGCTCGCCGC ATGACTGTG ATCGACACCA

FIG. 3K



pCMV-NS35

+2 N T L T C Y I K A R A A C R A A G L Q D C T M L V C G
6401 AACACCCCTCA CTTGCTACAT CAAGGCCCGG GCAGCCTGTC GAGCCGAGG GCTCCAGGAC TGCACCATGC TCGTGTGTGG
TTGTGGGAGT GAAGGATGTA GTTCCGGGCC CGTCGGACAG CTCGGGCTCC CGAGGTCTCTG ACGTGGTACG AGCACACACC

+2 D D L V V I C E S A G V Q E D A A S L R A F T E A M
6481 CGACGACTTA GTCGTTATCT GTGAAGCGC GGGGTCCAG GAGCAGCGG CGAGCCTGAG AGCCTTCACG GAGGCTATGA
GCTGCTGAAT CAGCAATAGA CACTTTCGG CCCCCAGGTC CTCCTGGGCC GCTCGGACTC TCGGAAGTGC CTCCGATACT

+2 T R Y S A P P G D P P Q P E Y D L E L I T S C S S N V
6561 CCAGGTACTC CGCCCCCCT GGGACCCCC CACAACCAGA ATAGGACTTG GAGCTCATAA CATCATGCTC CTCCAACGCTG
GGTCCATGAG GCGGGGGGA CCCCTGGGG GTGTTGCTCT TATGCTGAAC CTCGAGTATT GTAGTACGAG GAGGTTGCAC

+2 S V A H D G A G K R V Y Y L T R D P T T P L A R A A W
6641 TCAGTCGCCC ACGACGGCG TGGAAAGAGG GTCTACTACC TCACCCGTGA CCCTACAACC CCCCTCGGA GAGCTGCGTG
AGTCAGCGG TGCTGCCCGG ACCTTCTCC CAGATGATGG ACTGGCACT GGGATGTTGG GGGGAGCGCT CTCGACGCAC

+2 E T A R H T P V N S W L G N I I M F A P T L W A R M
6721 GGAGACAGCA AGACACATC CAGTCAATTC CTGGCTAGG AACATAATCA TGTTCGCCC CACACTGTGG GCGAGGATGA
CCTCTGCGT TCTGTGTGAG GTCAGTTAAG GACCGATCCG TTGTATTAGT ACAACGGGG GTGTGACACC CGCTCCTACT

+2 I L M T H F F S V L I A R D Q L E Q A L D C E I Y G A
6801 TACTGATGAC CCATTTCTTT AGCGTCCTTA TAGCCAGGA CCAGCTTGAA CAGGCCCTCG ATTGCGAGAT CTACGGGGCC
ATGACTACTG GGTAAGAAA TCGCAGGAAT ATCGGTCCT GTCCGGGAGC TAACGCTCTA GATGCCCCCG

+2 C Y S I E P L D L P P I I Q R L H G L S A F S L H S Y
6881 TGCTACTCCA TAGAACCACT GGATCTACCT CCAATCATC AAAGACTCCA TGGCCTCAGC GCATTTTCAC TCCACAGTTA
ACGATGAGGT ATCTTGGTGA CCTAGATGGA GGTTAGTAAG TTTCTGAGGT ACCGAGTCC CGTAAAGTG AGGTGTCAAT

FIG. 3L



pCMV-NS35

+2 S P G E I N R V A A C L R K L G V P P L R A W R H R
6961 CTCTCCAGGT GAAATCAATA GGTGGCGGC ATGCCTCAGA AAACCTGGGG TACCGCCCTT GCGAGCTTGG AGACACCGGG
GAGAGTCCA CTTAGTAT CCCACCGGG TAGGGAGTCT TTTGAACCC ATGGGGGAA CGCTCGAACC TCTGTGGCCC

+2 A R S V R A R L L A R G G R A A I C G K Y L F N W A V
7041 CCCGAGCGT CCGCGTAGG CTTCTGGCA GAGGAGGCAG GGCTGCCATA TGTGGCAAGT ACCTCTTCAA CTGGGCAGTA
GGCCTCGCA GCGCGGATCC GAAGACCGGT CTCCTCCGTC CCGACGGTAT ACACGGTTCA TGGAGAAGTT GACCCGTCAT

+2 R T K L K L T P I A A A G Q L D L S G W F T A G Y S G
7121 AGAACAAAGC TCAAACTCAC TCCAATAGCG GCGGCTGGCC AGCTGGACTT GTCCGGCTGG TTCACGGCTG GCTACAGCGG
TCTTGTTGG AGTTGAGTG AGTTATCGC CGGCGACCGG TCGACCTGAA CAGGCCGACC AAGTCCCGAC CGATGTCGCC

+2 G D I Y H S V S H A R P R W I W F C L L L L A A G V
7201 GGGAGACATT TATCACAGCG TGCTCATGCG CCGGCCCGCG TGGATCTGGT TTTGCCCTACT CCTGCTTGGT GCAGGGGTAG
CCCTCTGTAA ATAGTGTGCG ACAGAGTACG GGGCCGGGCG ACCTAGACCA AAACGGATGA GGACGAACGA CGTCCCCCATC

+2 G I Y L L P N R
7281 GCATCTACCT CCTCCCCAAC CGATGAAGGT TGGGGTAAAC ACTCCGGCCT AAAAAAATAA AAAATCTAG AAAGGCGCGC
CGTAGATGGA GGAGGGGTG GCTACTTCCA ACCCCATTG TGAGGCCGGA TTTTTTTT TTTTATGATC TTTCCGCGCG

BAMHI MluI

7361 CAAGATATCA AGGATCCACT ACGGTTAGA GCTCGCTGAT CAGCCTCGAC TGTGCTTCT AGTTGCCAGC CATCTGTTGT
GTTCTATAGT TCCTAGGTGA TCGGCAATCT CGAGCGACTA GTCCGAGCTG ACACGGAAGA TCAACGGTCTG GTAGACAACA

7441 TTGCCCCTCC CCCGTGCCTT CCTTGACCCT GGAAGGTGCC ACTCCCACTG TCCTTTCCTA ATAAAAATGAG GAAATTGCAT
AAGGGGAGG GGGCACGGAA GGAAGTGGGA CCTTCCACCG TGAGGGTGAC AGGAAAGGAT TATTTTACTC CTTTAACGTA

FIG. 3M



pCMV-NS35

7521	CGCATTGTCT GAGTAGGTGT CATTCTATTG TGGGGGTGG GGTGGGGCAG GACAGCAAGG GGGAGGATTG GGAAGACAAT CGGTAAACAGA CTCATCCACA GTAAGATAAG ACCCCCCACC CCACCCCGTC CTGTGTTCC CCTCCTAAC CCTTCTGTGA
7601	AGCAGGCATG CTGGGGAGCT CTTCCGGTTC CTGCTCACT GACTCGCTGC GCTCGGTGCT TCGGTGCGG CGAGCGGTAT TCGTCCGTAC GACCCCTCGA GAAGCGAAG GAGCGAGTGA CTGAGCGAGC CGAGCCAGCA AGCCAGCGC GCTCGCCATA
7681	CAGTCACTC AAAGGCGGTA ATACGGTTAT CCACAGAATC AGGGGATAAC GCAGGAAAGA ACATGTGAGC AAAAGGCCAG GTCGAGTGAG TTTCCGCCAT TATGCCAATA GGTGCTTAG TCCCTATTG CGTCTTTCT TGTACACTCG TTTTCCGGTC
7761	CAAAAGGCCA GGAACCGTAA AAAGGCGCG TTTGCTGGGT TTTTCCATAG GCTCCGCCCC CCTGACGAGC ATCACAATAA GTTTCCGGT CCTTGGCATT TTTCCGGCG AACGACCGCA AAAAGGTATC CGAGGCGGG GACTGCTCG TAGTGTTTT
7841	TCGACGCTCA AGTCAGAGGT GCGGAAACCC GACAGGACTA TAAAGATACC AGCGTTTCC CCTTGGAAAG TCCCTCGTGC AGCTGCGAGT TCAGTCTCCA CCGCTTTGG CTGTCTGAT ATTCTATGG TCCGAAAGG GGCACCTCG AGGAGCAGC
7921	GCTCTCCTGT TCCGACCCTG CCGCTTACCG GATACCTGTC CGCCTTTCTC CCTTCGGGA GCGTGGCGCT TTCTCAATGC CGAGAGGACA AGGCTGGAC GCGCAATGG CTATGGACAG GCGGAAAGAG GGAAGCCCTT CGCACCGCA AAGAGTTACG
8001	TCACGCTGTA GGTATCTCAG TTCGGTGTAG GTCTTCGCT CCAAGCTGGG CTGTGTGCAC GAACCCCCCG TTCAGCCCCA AGTGGACAT CCATAGAGTC AAGCCACATC CAGCAAGCGA GGTTCGACCC GACACACGTG CTGGGGGGC AAGTCGGGCT
8081	CCGCTGGCC TTATCCGGTA ACTATCGTCT TGAGTCCAAC CCGGTAAGAC ACGACTTATC GCCACTGGCA GCAGCCACTG GGCGACGGG AATAGGCCAT TGATAGCAGA ACTCAGTTG GGCATTCTG TGCTGAATAG CGTGACCGT CGTCGGTGAC
8161	GTAACAGGAT TAGCAGAGCG AGGTATGTAG GCGGTGCTAC AGAGTTCTTG AAGTGTGGC CTAACACTAGC CTACACTAGA CATTGTCTA ATCGTCTCG TCATACATC CGCCACGATG TCTCAAGAAC TTCACCCACG GATTGATGCC GATGTGATCT

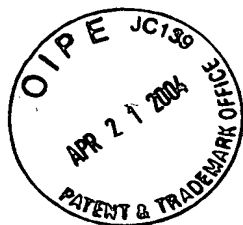
FIG. 3N



PCM-V-NS35

8241	AGGACAGTAT TTGGTATCTG CGCTCTGCTG AAGCCAGTTA CCTTCGGAAA AAGAGTTGGT AGCTCTTGAT CCGGCAAAACA TCCTGTCTATA AACCATAGAC GCGAGACGAC TTCGGTCAAT GGAAGCCTTT TTCTCAACCA TCGAGAACA GGCCTTTTGT
8321	AACCACCGCT GGTAGGGGTG GTTTTTTGT TTGCAAGCAG CAGATTACGC GCAGAAAAA AGGATCTCAA GAAGATCCTT TTGGTGGCGA CCATCGCCAC CAAAAAACA AAGTTCTGTC GTCTAATGCG CGTCTTTTTT TCCTAGAGTT CTCTAGGAA
8401	TGATCTTTTC TACGGGGTCT GACGTCAGT GGAACGAAAA CTCACGTTAA GGGATTTTGG TCATGAGATT ATCAAAAAAG ACTAGAAAAG ATGCCCCAGA CTGCGACTCA CCTTGTCTTT GAGTGCAATT CCTTAAACC AGTACTCTAA TAGTTTTTCC
8481	ATCTTCACCT AGATCCTTTT AAATTAATAA TGAAGTTTAA AATCAATCTA AAGTATATAT GAGTAAACTT GGCTGACAG TAGAAGTGGA TCTAGGAAAA TTTAATTTTT ACTTCAAAAT TTAGTTAGAT TTCAATATATA CTCATTGAA CCAGACTGTC
8561	TTACCAATGC TTAATCAGTG AGGCACCTAT CTCAGCGATC TGTCTATTTC GTTCATCCAT AGTTGCCCTGA CTCCTCCGCTG AATGTTTACG AATTAGTCA TCCGTGGATA GAGTCGCTAG ACAGATAAAG CAAGTAGGTA TCAACGGACT GAGGGGCAGC
8641	TGTAGATAAC TACGATACGG GAGGCTTAC CATCTGGCCC CAGTGTCTGA ATGATACCGC GAGACCCACG CTCACCGGCT ACATCTATTG ATGCTATGCC CTCCGGAATG GTAGACCGGG CTCACGACGT TACTATGGCG CTCTGGGTGC GAGTGGCCGA
8721	CCAGATTAT CAGCAATAA CCAGCCAGCC GGAAGGGCG AGCGCAGAG TGGTCTCTGA ACTTTATCCG CCTCCATCCA GGTCTAAATA GTCGTTATTT GGTCCGTCGG CCTTCCCGGC TCGCGTCTC ACCAGCACGT TGAAATAGGC GGAGGTAGGT
8801	GTCTATTAAT TGTGCGCGG AAGTAGAGT AAGTAGTTCG CCAGTTAATA GTTTGCGCAA CGTTGTTGCC ATTGCTACAG CAGATAATTA ACAACGGCCC TTCGATCTCA TTCAATCAAG GGTCAATAT CAACGCGTT GCAACAACCG TAACGATGTC
8881	GCATCTGTT GTACAGCTCG TCGTTTGGTA TGGCTTCATT CAGTCCGGT TCCCAACGAT CAAGGGGAGT TACATGATCC CGTAGCACCA CAGTGGGAG AGCAAAACAT ACCGAAGTAA GTCGAGGCCA AGGTTGCTA GTTCCGCTCA ATGTACTAGG

FIG. 30



pCMV-NS35

8961	CCCATGTTGT GCAAAAAGC GGTAGCTCC TTCGGTCCTC CGATCGTTGT CAGAAGTAAG TTGGCCGCGAG TGTATACACT GGGTACAACA CGTTTTTTCG CCAATCGAGG AAGCCAGGAG GCTAGCAACA GTCTTCATTC AACCGCGTC ACAATAGTGA
9041	CATGGTTATG GCAGCACTGC ATAAATCTCT TACTGTCTATG CCATCCGTA GATGCTTTTC TGTGACTGGT GAGTACTCAA GTACCAATAC CGTCGTGACG TATTAAGAGA ATGACAGTAC GGTAGGCAAT CTACGAAAAG AACTGACCA CTCATGAGTT
9121	CCAAGTCATT CTGAGAATAG TGTATGCGGC GACCGAGTTG CTCCTGCCCC GCGTCAATAC GGGATAATAC CGCGCCACAT GGTTCAGTAA GACTCTTATC ACATACGCGC CTGGCTCAAC GAGAACGGGC CGCAGTTATG CCTATTATG GCGCGGTGTA
9201	AGCAGAACTT TAAAAGTGCT CATCAATTGGA AAACGTTCTT CGGGGCGGAA ACTCTCAAGG ATCTTACCGC TGTGAGATC TCGCTTTGAA ATTTTCACGA GTAGTAACCT TTTGCAAGAA GCCCCGCTTT TGAGAGTTCC TAGAATGGCG ACAACTCTAG
9281	CAGTTCGATG TAACCCACTC GTGCACCCAA CTGATCTTCA GCATCTTTTA CTTTCACCAG CGTTTCTGGG TGAGCAAAA GTCAAGCTAC ATTGGGTGAG CACGTGGGTT GACTAGAAGT CGTAGAAAAT GAAAGTGGTC GCAAAGACCC ACTCGTTTTT
9361	CAGGAAGGCA AAATGCCGCA AAAAAGGGAA TAAGGGCGAC ACGGAATGT TGAATACTCA TACTCTTCCT TTTTCAATAT GTCCCTCCGT TTTACGGCGT TTTTTCCTT ATTCGCGCTG TGCCTTTACA ACTTATGAGT ATGAGAAGGA AAAAGTTATA
9441	TATTGAAGCA TTTATCAGG TTAATGTCTC ATGAGCGGAT ACATATTGA ATGTATTTAG AAAAATAAAC AAATAGGGGT ATAACTTCGT AAATAGTCCC AATAACAGAG TACTCGCCTA TGTATAAACT TACATAAATC TTTTATTG TTTATCCCCA
9521	TCCGCGCACA TTTCCCGGAA AAGTGCCACC TGACGTCTAA GAAACCATTA TTATCATGAC ATTAACCTAT AAAAATAGGC AGCGCGTGT AAAGGGGCTT TTCACGGTGG ACTGCAGATT CTTTGGTAAAT AATAGTACTG TAATTGGATA TTTTATCCG
9601	GTATCACGAG GCCCTTTCTG C CATAGTCTC CGGGAAGCA G

FIG. 3P

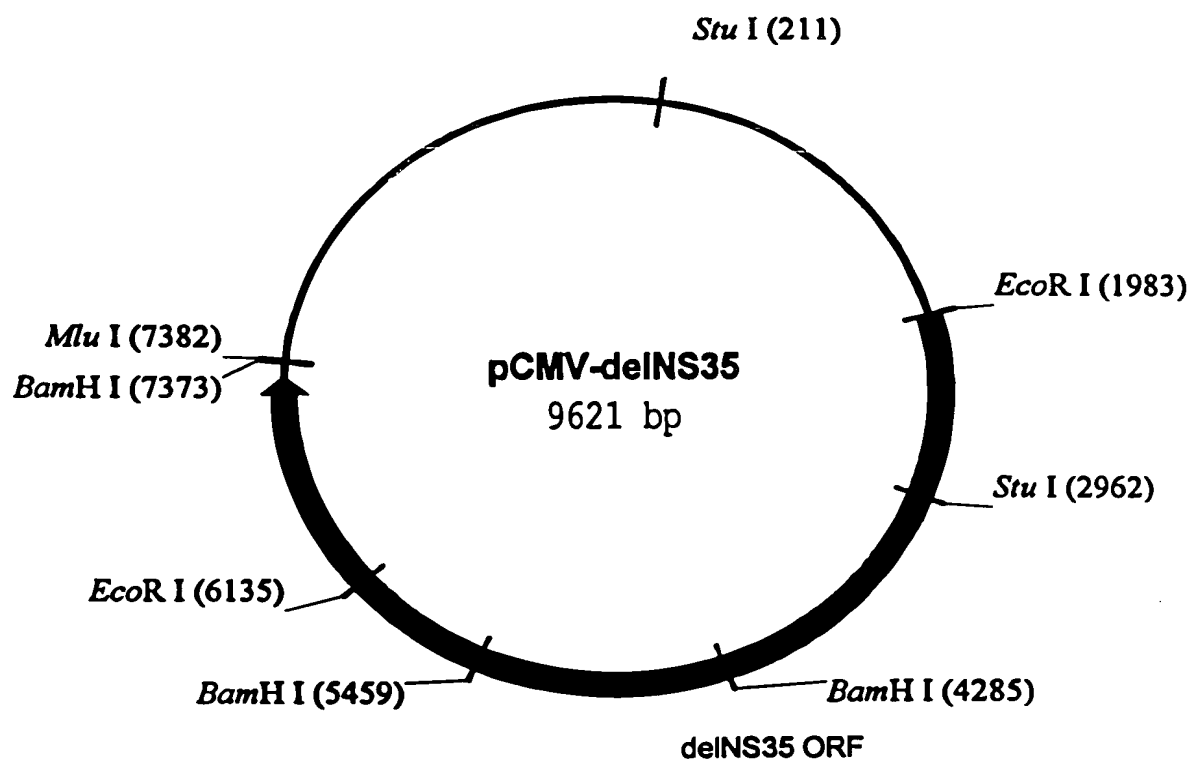


FIG . 4



pCMV-delNS35

1	TCGGCGGTTT CCGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG GAGACGGTCA CAGCTTGCT GTAAAGCGGAT AGCGCGCAAA CCCACTACTG CCACCTTTGG AGACTCTCTA CGTCGAGGCG CTCTGCCAGT GTCGAACAGA CATTCGCCTA
81	GCCGGGAGCA GACAAGCCCC TCAGGGCGCG TCAGGGGGTG TTGGCGGGTG TCGGGGCTGG CTAAACTATG CGGCATCAGA CGGCCCTCGT CTGTTGGGGC AGTCCGGCGC AGTCGCCAC AACCGCCAC AGCCCGGACC GAATTGATAC GCCGTAGTCT
161	GCAGATTCTA CTGAGAGTGC ACCATATGAA GCTTTTGGCA AAAGCCTAGG CCTCCAAAA AGCCTCCTCA CTACTTCTGG CGTCTAACAT GACTCTCAG TGGTATACTT CGAAAAAGCT TTTCGGATCC GGAGGTTTT TCGGAGGAGT GATGAAGACC
241	AATAGCTCAG AGGCGGAGGC GGCCTCGGC TCTGCATAAA TAAAAAAAT TAGTCAGCCA TGGGGCGGAG ATTGGGCGGA TTATCGAGTC TCGGCTCCG CCGGAGCGG AGACGTATT ATTTTTTTTA ATCAGTCGGT ACCCGCCTC TTACCCGCCT
321	ACTGGGCGG GAGGAATTA TTGGCTATT GCCATTGCAT ACGTTGTATC TATATCATAA TATGTACATT TATATTGGCT TGACCCGCCC CTCCTTAAT AACCGATAAC CGGTAACTG TGCAACATAG ATATAGTATT ATACATGTAA ATATAACCGA
401	CATGTCCAAT ATGACCGCCA TGTTGACATT GATTATTGAC TAGTTATTAA TAGTAATCAA TTACGGGGTC ATTAGTTCAT GTACAGTTA TACTGGGGT ACAACTGTAA CTAAATACTG ATCAATAATT ATCATTAGTT AATGCCCCAG TAATCAAGTA
481	AGCCCATATA TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAAGGACC CCGCCCCATT TCGGGTATAT ACCTCAAGGC GCAATGTATT GAATGCCATT TACCGGGCGG ACCGACTGGC GGGTTGCTGG GGGCGGGTAA
561	GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTACGGT CTGCAGTTAT TACTGCATAC AAGGTATCA TTGCGGTAT CCCTGAAAGG TAACTGCAGT TACCCAGCTC ATAAATGCCA

StuI

FIG. 5A



pCMV-delINS35

641 AAACGTGCCC CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTCCGGCC CCTATTGACG TCAATGACGG TAAATGGCCC
TTTGACGGGT GAACCGTTCAT GTAGTTCACA TAGTATACGG TTCAGGGGGG GGATAACTGC AGTTACTGCC ATTTACCGGG

721 GCCTGGCATT ATGCCCAGTA CATGACCTTA CGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA TCGCTATTAC
CGGACCGTAA TACGGGTTCAT GTACTGGAAT GGCCTGAAAG GATGAACCGT CATGTAGATG CATAATCAGT ACGGATAATG

801 CATGGTGATG CGGTTTGGC AGTACACCAA TGGGCGTGGA TAGCGGTTTG ACTCACGGGG ATTTCCAAAGT CTCCACCCCA
GTACCACTAC GCCAAACCG TCAATGTGTT ACCCGCACCT ATGCCCAAAC TGAGTGGCCC TAAAGTTCA GAGTGGGGT

881 TTGACGTCAA TGGGAGTTTG TTTTGGCACC AAAATCAACG GGACTTTCCA AAATGTGTA ATAAACCCCG CCCGTTGACG
AACTGCAGTT ACCCTCAAAC AAAACCGTGG TTTTAGTTGC CCTGAAAGGT TTTACAGCAT TATTGGGGCG GGGCAACTGC

961 CAAATGGCG GTAGCGGTGT ACGGTGGAG GTCTATATAA GCAGAGCTCG TTTAGTGAAC CGTCAGATCG CCTGGAGACG
GTTTACCGC CATCCGCACA TGCCACCTC CAGATATATT CGTCTCGAGC AAATCACTTG GCAGTCTAGC GGACCTCTGC

1041 CCATCCACGC TGTTTTGACC TCCATAGAAG ACACCGGGAC CGATCCAGCC TCCGGGGCGG GGAACGGTGC ATTGGAACGC
GGTAGGTGG ACAAAACTGG AGGTATCTTC TGTGGCCCTG GCTAGGTGCG AGGCGCCCGC CCTTGCCACG TAACCTTGCG

1121 GGATTCCCC TGCCAAGAGT GACGTAACTA CCGCCTATAG ACTCTATAGG CACACCCCTT TGGCTCTTAT GCATGCTATA
CCTAAGGGC ACGGTTCTCA CTGCATTTCAT GCGGGATATC TGAGATATCC GTGTGGGAA ACCGAGAATA CGTACGATAT

1201 CTGTTTTTGG CTTGGGGCCT ATACACCCCG CCTCCTTATG CTATAGGTGA TGGTATAGCT TAGCCTATAG GTGTGGGTTA
GACAAAAAC GAACCCCGA TATGTGGGG CGAGGAATAC GATATCCACT ACCATATCGA ATCGGATATC CACACCCAAT

1281 TTGACCATT TTAGCCACTC CCCTATTGCT GACGATACTT TCCATTACTA ATCCATAACA TGGCTCTTTG CCACAACCTAT
AACTGCTAAT AACTGGTGAG GGGATAACCA CTGCTATGAA AGGTAATGAT TAGGTATTGT ACCGAGAAAC GGTGTTGATA

FIG. 5B



pCMV-delINS35

1361	CTCTATTGGC TATATGCCAA TACTCTGTCC TTCAGAGACT GACACGGACT CTGTATTTT ACAGGATGGG GTCCATTAT	
	GAGATAACCG ATATACGGTT ATGAGACAGG AAGTCTCTGA CTGTGCTGA GACATAAAAA TGTCCTACCC CAGGTAAATA	
1441	TATTTAACAA TTCACATATA CAACAACGCC GTCCCCCGTG CCGGCAGTTT TTATTAAACA TAGCGTGGGA TCTCCGACAT	
	ATAAATGTTT AAGTGATAT GTTGTTCGG GAGGGGGCAC GGGGTCAAA AATAATTTCT ATCGCACCTT AGAGGCTGTA	
1521	CTCGGGTAGG TGTTCGGAC ATGGGCTCTT CTCCGGTAGC GCGGGAGCTT CCACATCCGA CCCCTGGTCC CATCCGTCCA	
	GAGCCCATGC ACAAGGCTG TACCCGAGAA GAGGCCATCG CCGCCTCGAA GGTGTAGGCT CGGGACCAGG GTAGGCAGGT	
1601	GCGGCTCATG GTCGCTCGG AGCTCCTTGC TCCTAACAGT GGAGGCCAGA CTTAGGACA GCACAATGCC CACCACCACC	
	CGCCGAGTAC CAGCGAGCCG TCGAGGAACG AGGATTGTA CCTCGGTCT GAATCCGTCT CGTGTIACGG GTGGTGGTGG	
1681	AGTGTGCGG ACAAGGCCGT GCGGGTAGG TATGTGTCTG AAAATGAGCT CGGAGATTGG GCTCGCACCT GGACGCAGAT	
	TCACACGGCG GTTTCGGCA CCGCCATCCC ATACACAGAC TTTTACTCGA GCCTCTAACC CGAGCGTGA CCTGGCTCTA	
1761	GGAAAGCTTA AGGCAGCGG AGAAGAAGAT GCAGGCAGCT GAGTTGTTGT ATTCTGATAA GAGTCAGAGG TAACITCCCGT	
	CCITTCGTAAT TCCGTGCGG TCTTCTTCTA CGTCCGTCCA CTCAACAACA TAAGACTATT CTCAGTCTCC ATTGAGGGCA	
1841	TGCGGTGCTG TTAACGGTGG AGGGCAGTGT AGTCTGAGCA GTACTCGTGT CTGCGCGCGG CGCCACCAGA CATAATAGCT	
	ACGCCACGAC AATTGCCACC TCCCGTCACA TCAGACTCGT CATGAGCAAC GACGGCGCGG CCGGTGGTCT GTATTATCGA	
+2		M A A
		EcoRI

1921	GACAGACTAA CAGACTGTTT CTTTCCATGG GTCTTTTCTG CAGTCACCGT CGTCGACCTA AGAATTCACC ATGGCTGCAT	
	CTGTCTGATT GTCTGACAA GAAAGGTACC CAGAAAGAC GTCAGTGGCA GCAGCTGGAT TCTTAAGTGG TACCGACGTA	

FIG. 5C



pCMV-delNS35

+2 Y A A Q G Y K V L V L N P S V A A T L G F G A Y M S K
2001 ATGCAGCTCA GGGCTATAAG GTGCTAGTAC TCAACCCCTC TGTTGCTGCA ACACCTGGCT TTGGTGCTTA CATGTCCAAG
TAGCTCGAGT CCCGATATTC CACGATCATG AGTTGGGAG ACAACGACCT TGTGACCCGA AACCAAGCAAT GTACAGGTTG

+2 A H G I D P N I R T G V R T I T T G S P I T Y S T Y G
2081 GCTCATGGGA TCGATCCTAA CATCAGGACC GGGGTGAGAA CAATTACAC TGGCAGCCCC ATCAGGTACT CCACCTACGG
CGAGTACCCT AGCTAGGATT GTAGTCCTGG CCCCACTCTT GTTAATGCTG ACCGTGGGG TAGTGCATGA GGTGGATGCC

+2 K F L A D G G C S G G A Y D I I I C D E C H S T D A
2161 CAAGTTCCTT GCCGACGGCG GTGCTCGGG GGGCGCTTAT GACATAATAA TTGTGACGA GTGCCACTCC ACGGATGCCA
GTTCAAGGAA CGGCTGCCG CCACGAGCCC CCGGGAATA CTGTATTATT AAACACTGCT CACGGTGAGG TGCCTACGGT

+2 T S I L G I G T V L D Q A E T A G A R L V V L A T A T
2241 CATCCATCTT GGGCATTCG ACTGTCTTG ACCAAGCAGA GACTCGGGG GCGAGACTGG TTGTGCTCGC CACCGGCACC
GTAGGTAGAA CCGGTAACCG TGACAGGAAC TGTTCTGCT CTGACGCCCC CGCTCTGACC AACACGAGCG GTGGCGGTGG

+2 P P G S V T V P H P N I E E V A L S T T G E I P F Y G
2321 CCTCCGGCT CCGTCACTGT GCCCATCCC AACATCGAGG AGTTGCTCT GTCCACCACC GGAGAGATCC CTTTITACGG
GGAGGCCCGA GGCAGTGACA CCGGGTAGGG TTGTAGCTCC TCCAACGAGA CAGGTGGTGG CCTCTCTAGG GAAAAATGCC

+2 K A I P L E V I K G G R H L I F C H S K K K C D E L
2401 CAAGGCTATC CCCCTCGAAG TAATCAAGGG GGGGAGACAT CTCATCTTCT GTCATTCAAA GAAGAAGTGC GACGAATCG
GTTCCGATAG GGGGAGCTTC ATTAGTCCC CCCCTCTGTA GAGTAGAAGA CAGTAAGTTT CTTCTTCACG CTGCTTGAGC

+2 A A K L V A L G I N A V A Y Y R G L D V S V I P T S G
2481 CCGCAAAGCT GGTCCGATTG GGCATCAATG CCGTGGCCTA CTACCGGCT CTTGACGTGT CCGTCATCCC GACCAAGCGG
GGCGTTTCCA CCAGCGTAAC CCGTAGTTAC GGCACGGAT GATGGGACA GAACTGCACA GGCAGTAGG CTGCTCGCCG

FIG. 5D



pCMV-deINS35

+2 D V V V V A T D A L M T G Y T G D F D S V I D C N T C
2561 GATGTGTGG TCGTGGCAAC CGATGCCCTC ATGACCGGT ATACCGGCGA CTTCGACTCG GTGATAGACT GCAATACGTG
CTACAACAGC AGCACCGTTG GCTACGGGAG TACTGGCCGA TATGGCCGCT GAAGCTGAGC CACTATCTGA CGTTATGCAC

+2 V T Q T V D F S L D P T F T I E T I T L P Q D A V S
2641 TGTACCCAG ACAGTCGATT TCAGCCTTGA CCTACCTTC ACCATTGAGA CAATCAGCT CCCCCAAGAT GCTGTCTCC
ACAGTGGGTC TGTACGCTAA AGTCGGAACT GGGATGGAAG TGGTAACCTCT GTTAGTCCGA GGGGGTTCTA CGACAGAGGG

+2 R T Q R R G R T G R G K P G I Y R F V A P G E R P S G
2721 GCACTCAACG TCGGGGCAGG ACTGGCAGG GGAAGCCAGG CATCTACAGA TTGTGGCAC CGGGGGAGCG CCCCTCCGGC
CGTGAGTTGC AGCCCCGTCC TGACCGTCCC CCTTCGGTCC GTAGATGTCT AAACACCGTG CCCCCCTCGC GGGGAGGGCG

+2 M F D S S V L C E C Y D A G C A W Y E L T P A E T T V
2801 ATGTTGCACT CGTCCGTCTT CTGTGAGTGC TATGACGCG GCTCTGCTTG GTATGAGCTC ACGCCGCGCG AGACTACAGT
TACAAGCTGA GCAGGCAGGA GACACTCAGG ATACTGCGTC CGACACGAA CATACTCGAG TCGGGGGCG TCTGATCTCA

+2 R L R A Y M N T P G L P V C Q D H L E F W E G V F T
StuI

2881 TAGGCTACGA GCGTACATGA ACACCCCGGG GCTTCCCGTG TGCCAGGACC ATCTTGAATT TTGGGAGGGC GTCTTTACAG
ATCCGATGCT CGCATGTACT TGTGGGGCCC CGAAGGGCAC ACGTCTCTGG TAGAACTTAA AACCTCCCG CAGAAATGTC

+2 G L T H I D A H F L S Q T K Q S G E N L P Y L V A Y Q
StuI

2961 GCCTCACTCA TATAGATGCC CACTTTCTAT CCCAGACAAA GCAGAGTGGG GAGAACCTTC CTTACCTGGT AGGCTACCAA
CGGAGTGAGT ATATCTACGG GTGAAAGATA GGGTCTGTTT CGTCTCACCC CTCCTTGAAG GAATGGACCA TCGCATGGTT

FIG. 5E



pCMV-delINS35

+2 A T V C A R A Q A P P P S W D Q M W K C L I R L K P T
3041 GCCACCGTGT GCGCTAGGC TCAAGCCCTT CCCCCTCGT GGCACCATGT GTGGAAGTGT TTGATTGGC TCAAGCCAC
CGGTGGCACA CGCGATCCG AGTTCCGGGA GGGGGTAGCA CCTGGTCTA CACCTTCACA AACTAAGCG AGTTCGGGTG

+2 L H G P T P L L Y R L G A V Q N E I T L T H P V T K
3121 CCTCCATGGG CCAACACCC TGCTATACAG ACTGGCGCT GTTCAGAATG AATCACCTT GACGCACCCA GTCACCAAT
GGAGGTACCC GGTGTGGG ACATATGTC TGACCCCGGA CAAGTCTTAC TTAGTGGGA CTGCGTGGT CAGTGGTTT

+2 Y I M T C M S A D L E V V T S T W V L V G G V L A A L
3201 ACATCATGAC ATGCATGTC GCGACCTGG AGTCTGTAC GACACCTGG GTGCTCGTTG GCGCGCTCTT GGCTGCTTTG
TGATGACTG TACGTACAG CGGCTGGACC TCCAGCAGT TCCGTGGACC CACGAGCAAC CGCCGACGA CCGACGAAC

+2 A A Y C L S T G C V V I V G R V V L S G K P A I I P D
3281 GCGCGTATT GCCTGTCAAC AGGCTGCGT GTCATAGTG GCAGGTCTGT CTGTGCGG AAGCGGCAA TCATACCTGA
CGCGGCATAA CGGACACTTG TCCGACGCAC CAGTATCACC CGTCCACGA GAACAGGCC TTGCGCGTT AGTATGGACT

+2 R E V L Y R E F D E M E E C S Q H L P Y I E Q G M M
3361 CAGGGAAGTC CTCTACGAG AGTTCGATGA GATGGAAG TGCTCTCAGC ACTTACCGTA CATCGAGCAA GGGATGATGC
GTCCCTTCAG GAGATGGTC TCAAGCTACT CTACCTTCTC ACCAGAGTCG TGAATGGCAT GTAGCTCGTT CCTACTACG

+2 L A E Q F K Q K A L G L L Q T A S R Q A E V I A P A V
3441 TCGCCGAGCA GTTCAAGCAG AAGCCCTCG GCCTCCTGCA GACCGGTCC CGTCAGGCAG AGTTATCGC CCTGCTGTC
AGCGGCTCGT CAAGTTCGT TTCCGGGAGC CGGAGGACGT CTGGCGCAGG GCAGTCCGTC TCCAATAGCG GGGACGACAG

+2 Q T N W Q K L E T F W A K H M W N F I S G I Q Y L A G
3521 CAGACCAACT GGCATAAACT CGAGACCTTC TGGCGGAGC ATATGTGGA CTTCATCAGT GGGATACAAT ACTTGGCGGG
GTCTGGTTGA CCGTTTTTGA GCTCTGGAAG ACCCGCTTGG TATACACCTT GAAGTAGTCA CCTATGTTA TGAACCGCGC

FIG. 5F



pCMV-delINS35

+2 L S T L P G N P A I A S L M A F T A A V T S P L T T
3601 CTGTCAACG CTGCTGGTA ACCCGGCAT TGCTTCATTG ATGGCTTTTA CAGCTGCTGT CACCAGCCCA CTAACCACTA
GAACAGTTGC GACGGACCAT TGGGGCGGTA AGGAAGTAAC TACCGAAAT GTGGTCGGGT GATTGCTGAT

+2 S Q T L L F N I L G G W V A A Q L A A P G A A T A F V
3681 GCCAAACCTT CCTCTTCAAC ATATTGGGG GGTGGTGGC TGCCAGCTC GCCGCCCCG GTGCGGTAC TGCCTTTCTG
CGCTTGGGA GGAGAAGTTG TATAACCCC CCACCCACCG ACGGTGCGAG CGGCGGGGGC CACGGCGATG ACGGAACAC

+2 G A G L A G A A I G S V G L G K V L I D I L A G Y G A
3761 GCGCTGGCT TAGCTGGCG CGCATGGC AGTGTGGAC TGGGAAGGT CCTCATAGAC ATCCTTGCAG GGTATGGCG
CGCGACCGA ATCGACCGG GCGTAGCGG TCACAACCTG ACCCTTCCA GGAGTATCTG TAGGAACGTC CCATACCGG

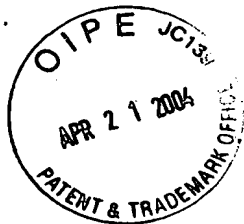
+2 G V A G A L V A F K I M S G E V P S T E D L V N L L
3841 GCGCTGGCG GGAGCTCTT TGGCATTCAA GATCATGAGC GGTGAGTCC CCTCCACGGA GGACCTGTC AATCTACTGC
CCCGCACCGC CCTCGAGAAC ACGTAAGTT CTAGTACTCG CCACTCCAGG GGAGTGCCT CCTGGACCAG TTAGATGACG

+2 P A I L S P G A L V V G V V C A A I L R R H V G P G E
3921 CCGCATCCT CTGCGCCGGA GCGCTCGTAG TCGGCGTGGT CTGTGCACCA ATACTCGCC GGCACGTTGG CCGGGCGGAG
GCGGTAGGA GACCGGGCCT CCGGAGCATC AGCCGCACCA GACACGTCGT TATGACGCGG CCGTGCAACC GGGCCGCTC

+2 G A V Q W M N R L I A F A S R G N H V S P T H Y V P E
4001 GGGCAGTGC AGTGGATGAA CCGCTGATA GCCTTGGCT CCGGGGGAA CCATGTTTCC CCGACGCACT ACGTCCCGGA
CCCCGTACG TCACCTACTT GCGGACTAT CGGAAGCGGA GGGCCCTT GGTACAAAGG GGTGCGTGA TGCACGGCCT

+2 S D A A A R V T A I L S S L T V T Q L L R R L H Q W
4081 GAGCGATGA CTTGCGCGG TCACTGCCAT ACTCAGCAGC CTCACGTAA CCCAGCTCCT GAGCGACTG CACCACTGGA
CTCGTACGT CGACGGGCG AGTGACGGTA TGAGTGTGCG GAGTACATT GGTGCGAGGA CTCCGCTGAC GTGGTCACT

FIG. 5G



pCMV-deINS35

+2 I S S E C T T P C S G S W L R D I W D W I C E V L S D
4161 TAAGCTCGA GTGTACCACT CCATGCTCCG GTTCTGGCT AAGGACATC TGGGACTGGA TATCGGAGGT GTTGAGCGAC
ATTGAGCCT CACATGGTGA GGTACGAGGC CAAGGACCGA TTCCCTGTAG ACCCTGACCT ATAGGTCCA CAACTCGCTG

+2 F K T W L K A K L M P Q L P G I P F V S C Q R G Y K G
BamHI

4241 TTTAAGACT GGCTAAAAGC TAAGCTCATG CCACAGCTGC CTGGGATCCC CTTTGTGTCC TGCCAGCCGC GGTATAAGGG
AAATTCTGA CCGATTTTCG ATTGAGTAC GGTGTGACG GACCTAGG GAAACACAGG ACGGTGCGGC CCATATTCCC

+2 V W R G D G I M H T R C H C G A E I T G H V K N G T
4321 GGTCTGGCA GGGACGGCA TCATGCACAC TCGTGGCAG TGTGGAGCTG AGATCACTGG ACATGTCAA AACGGGACGA
CCAGACCGCT CCCCTGCCGT AGTACGTGTG ACGGACGGTG ACACCTGCAC TCTAGTGACC TGTACAGTTT TTGCCCTGCT

+2 M R I V G P R T C R N M W S G T F P I N A Y T T G P C
4401 TCAGGATCGT CGGTCCCTAGG ACCTGCAGGA ACATGTGGAG TGGACCTTC CCCATTATG CCTACACCAC GGGCCCCCTGT
ACTCCTAGCA GCCAGGATCC TGGACGTCTT TGTACACCTC ACCCTGGAAG GGTAAATTAC GGATGTGGTG CCGGGGGACA

+2 T P L P A P N Y T F A L W R V S A E E Y V E I R Q V G
4481 ACCCCCTTC CTGCGCCGAA CTACACCTTC GCGCTATGGA GGGTGTCTGC AGAGGAATAC GTGGAGATAA GGCAGGTGGG
TGGGGGAAG GACGGGCTT GATGTGCAAG CGGATACCT CCCACAGAGC TCTCCTATG CACCTCTATT CCGTCCACCC

+2 D F H Y V T G M T T D N L K C P C Q V P S P E F F T
4561 GGACTTCCAC TACGTGACGG GTATGACTAC TGACAATCTT AAATGCCGT GCCAGGTCCC ATGCCCCGAA TTTTTCACAG
CCTGAAGTG ATGCACCTGCC CATACTGATG ACTGTTAGAA TTTACGGGCA CCGTCCAGG TAGCGGGCTT AAAAAGTGC

+2 E L D G V R L H R F A P P C K P L L R E E V S F R V G
4641 AATTGACGG GGTGCGCCTA CATAGGTTTG CGCCCCCTG CAAGCCCTTG CTGCGGGAGG AGGTATCATT CAGAGTAGGA
TTAACCTGCC CCACGGGAT GTATCCAAAC GCGGGGGGAC GTTCGGGAAC GACGCCCTCC TCCATAGTAA GTCTCATCCT

FIG. 5H



pCMV-deINS35

+2 L H E Y P V G S Q L P C E P E P D V A V L T S M L T D
4721 CTCACGAAT ACCGGTAGG GTCCGAATTA CCTTCGAGC CCGAACCGGA CCGTGGCCGTG TTGACGTCCA TGCTCACTGA
GAGGTGCTTA TGGGCCATCC CAGCGTTAAT GGAACGCTCG GCGTTGGCCT GCACCGGCAC AACTGCAGGT ACGAGTGACT

+2 P S H I T A E A A G R R L A R G S P P S V A S S S A
4801 TCCCTCCCAT ATAACAGCAG AGGGGCGCG CCGAAGGTTG GCGAGGGGAT CACCCCCCTC TGTGGCCAGC TCCTCGGCTA
AGGAGGGTA TATTGCTC TCCGCGGCG CCGTTCCAAC CCGTCCCTA GTGGGGGAG ACACCGGTG AGGAGCCGAT

+2 S Q L S A P S L K A T C T A N H D S P D A E L I E A N
4881 GCCAGCTATC CGCTCCATCT CTCAGGCA CTTGCACCGC TAACCATGAC TCCCTGATG CTGAGCTCAT AGAGGCCAAC
CGCTCGATAG GCGAGGTAGA GAGTCCGTT GAACGTGGCG ATTGTA CTG AGGGACTAC GACTCGAGTA TCTCCGTTG

+2 L L W R Q E M G G N I T R V E S E N K V V I L D S F D
4961 CTCCTATGA GCGAGGAGAT GCGGGCAAC ATCACCAGG TTGAGTCAGA AAACAAAGTG GTGATTCTGG ACTCCTTCCA
GAGATACCT CCGTCTCTA CCGCGGTTG TAGTGTCCC AACTCAGTCT TTTGTTTTCAC CACTAAGACC TGAGGAAGCT

+2 P L V A E E D E R E I S V P A E I L R K S R R F A Q
5041 TCCGTTGTG GCGGAGGAG ACGAGGCGGA GATCTCCGA CCGCAGAAA TCCTGCGGAA GTCTCGGAGA TTCGGCCAGG
AGCGAACAC CGCTCCTCC TGCTGCGCCT CTAGAGGCAT GGGCGTCTT AGGACGCCCT CAGAGCCTCT AAGCGGCTCC

+2 A L P V W A R P D Y N P P L V E T W K K P D Y E P P V
5121 CCTTCCCGT TTGGGCGCG CCGACTATA ACCCCCGCT AGTGGAGAG CCGGAGGAG CCGACTACGA ACCACTGTG
GGGAGGGCA AACCGCGCG GCGCTGAT TGGGGGCGA TCACCTCTG ACCTTTTCG GCGTGATGCT TGGTGACAC

+2 V H G C P L P P P K S P P V P P P R K K R T V V L T E
5201 GTCCATGGCT GCGCGCTCC ACCTCAAAAG TCCCTCCTG TGCCTCCGC TCGAAGGAG CCGAGCGGTG TCCTCACTGA
CAGTACCGA CCGGCGAAG TGGAGTTTC AGGGAGGAG ACGGAGGCG AGCTTCTTC GCCTGCCACC AGGAGTGACT

FIG. 5I



pCMV-deINS35

+2 S T L S T A L A E L A T R S F G S S S T S G I T G D
5281 ATCAACCCTA TCTACTGCCT TGGCCGAGCT CGCCACAGCA AGCTTTGGCA GCTCCTCAAC TTCCGGCATT ACGGGGGACA
TAGTTGGGAT AGATGACGGA ACCGGCTCGA GCGGTGCTCT TCGAAACCGT CGAGGAGTTG AAGGCGTAA TGCCCGCTGT

+2 N T T T S S E P A P S G C P P D S D A E S Y S S M P P
5361 ATACGACAAC ATCCTCTGAG CCGGCCCTTT CTGGCTGCC CCGGACTCC GACGCTGAGT CCTATTCTC CATGCCCCCC
TATGCTGTTG TAGGAGACTC GGGCGGGGAA GACCGACGG GGGGCTGAGG CTGCGACTCA GGATAAGGAG GTACGGGGGG

+2 L E G E P G D P D L S D G S W S T V S S E A N A E D V
BamHI

5441 CTGGAGGGG AGCCTGCGCA TCCGGATCTT ACGGACGGGT CATGGTCAAC GGTCAGTAGT GAGGCCAACG CGGAGGATGT
GACCTCCCC TCGGACCCCT AGCCCTAGAA TCGCTGCCCC GTACCAGTTG CCAGTCATCA CTCGGGTTGC GCCTCCTACA

+2 V C C S M S Y S W T G A L V T P C A A E E Q K L P I
5521 CGTGCTGTC TCAATGCTT ACTCTGGAC AGGCGCACTC GTCACCCCGT GCGCGCGGA AGAACAGAAA CTGCCCATCA
GCACACGAG AGTIACAGAA TGAGAACCTG TCCGGGTGAG CAGTGGGCA CCGCGCGCT TCTGTCTTT GACGGGTAGT

+2 N A L S N S L L R H H N L V Y S T T S R S A C Q R Q K
5601 ATGCACTAAG CAACTCGTTG CTACGTACC ACAATTGGT GTATTCCACC ACCTCACGA GTGCTTGCCA AAGGCAGAAG
TACGTGATTG GTGAGCAAC GATGCAGTGG TGTTAAACCA CATAAGGTGG TGGAGTGGT CACGAACGGT TTCCGTCTTC

+2 K V T F D R L Q V L D S H Y Q D V L K E V K A A S K
5681 AAAGTCACAT TTGACAGACT GCAAGTTCTG GACAGCCATT ACCAGGACGT ACTCAAGGAG GTTAAAGCAG CGGCGTCAAA
TTTCAGTGA AACTGTCTGA CGTCAAGAC CTGTCGGTAA TGGTCTGCA TGAGTCTCTC CAATTCTGTC GCGGCAGTTT

+2 V K A N L L S V E E A C S L T P P H S A K S K F G Y
5761 AGTGAAGGCT AACTTGCTAT CCGTAGAGGA AGCTTGCAGC CTGACGCCCC CACACTCAGC CAAATCCAAG TTTGGTTATG
TCACCTCCGA TTGAACGATA GGCATCTCCT TCGAACGTCG GACTGCGGGG GTGTGAGTCG GTTTAGGTTT AAACCAATAC

FIG. 5J



pCMV-delNS35

+2 G A K D V R C H A R K A V T H I N S V W K D L L E D N
5841 GGGCAAAAGA CGTCCGTTGC CATGCCAGAA AGGCCGTAAC CCACATCAAC TCCGTGTGGA AAGACCTTCT GGAAGACAAT
CCCCTTTTCT GCAGGCAACG GTACGGTCTT TCCGGCATTG GGTGTAGTGG AGGCACACCT TTCTTGAAGA CCTTCTGTTA

+2 V T P I D T T I M A K N E V F C V Q P E K G G R K P A
5921 GTAAACACAA TAGACACTAC CATCATGGCT AAGAACGAGG TTTTCTGCGT TCAGCCTGAG AAGGGGGTC GTAAGCCAGC
CATTCGTGTT ATCTGTGATG GTAGTACCGA TTCTTGCTCC AAAAGACGGA AGTCGGAATC TTCCCCCAG CATTCGCTCG

+2 R L I V F P D L G V R V C E K M A L Y D V V T K L P
6001 TCGTCTCATC GTGTTCCCG ATCTGGGCGT GCGCGTGTGC GAAAAGATGG CTTTGTACGA CGTGGTTACA AAGCTCCCT
AGCAGAGTAG CACAAGGGG TAGACCGCA GCGGCACAG CTTTTCTACC GAAACATGCT GCACCAATGT TTCGAGGGGA

+2 L A V M G S S Y G F Q Y S P G Q R V E F L V Q A W K S
EcoRI

6081 TGGCGTGAT GGAAGCTCC TACGGATTCC AATACTACC AGGACAGCGG GTTGAATTCC TCGTGCAAGC GTGGAAGTCC
ACCGGCACTA CCCTTCGAGG ATGCCCTAAGG TTATGAGTGG TCCTGTGCGC CAACTTAAGG AGCAGGTTCC CACCTTCAGG

+2 K K T P M G F S Y D T R C F D S T V T E S D I R T E E
6161 AAGAAAACCC CAATGGGGTT CTCGTATGAT ACCCGCTGCT TTGACTCCAC AGTCACTGAG AGGCACATCC GTACGGAGGA
TTCTTTTGGG GTTACCCCAA GAGCATACTA TGGGGGACGA AACTGAGGTG TCACTGACTC TCGCTGTAGG CATGCCCTCCT

+2 A I Y Q C C D L D P Q A R V A I K S L T E R L Y V G
6241 GGCAATCTAC CAATGTTGT ACCTCGACCC CCAAGCCCGC GTGGCCATCA AGTCCCTCAC CGAGAGGCTT TATGTTGGGG
CCGTTAGATG GTTACAACAC TGGAGTGGG GGTTCGGGGC CACCGGTAGT TCAGGGAGTG GCTCTCGAA ATACAACCC

+2 G P L T N S R G E N C G Y R R C R A S G V L T T S C G
6321 GCCCTCTTAC CAATTCAAGG GGGGAGAACT GCGGTATCG CAGGTGCGC GCGAGCGGCG TACTGACAAC TAGCTGTGGT
CGGAGAAATG GTTAAGTTCC CCCCTTTGA CGCCGATAGC GTCCACGGCG CGCTCGCCCG ATGACTGTTG ATCGACACCA

FIG. 5K



pCMV-delNS35

+2 N T L T C Y I K A R A A C R A A G L Q D C T M L V C G
6401 AACACCTCA CTGCTACAT CAAGGCCGG GCAGCTGTC GAGCGCAGG GCTCCAGGAC TGCACCATGC TCGTGTGTGG
TTGTGGGAGT GAACGATGTA GTTCCGGGCC CGTCGGACAG CTCGGCGTCC CGAGGTCCTG ACCTGGTAGG AGCACACACC

+2 D D L V V I C E S A G V Q E D A A S L R A F T E A M
6481 CGACGACTTA GTCGTTATCT GTGAAAGCG GGGGTCCAG GAGGACGGG CGAGCCTGAG AGCCTTCACG GAGGCTATGA
GCTGCTGAAT CAGCAATAGA CACTTTCGGC CCCCAGGTC CTCCTGGCC GCTCGGACTC TCGGAAGTGC CTCCGATACT

+2 T R Y S A P P G D P P Q P E Y D L E L I T S C S S N V
6561 CCAGGTACTC CGCCCCCCT GGGGACCCCC CACAACCAGA ATACGACTTG GAGCTCATAA CATCATGCTC CTCCAACGCTG
GGTCCATGAG CGGGGGGGA CCCCTGGGG GTGTGGTCT TATGCTGAAC CTCGAGTATT GTAGTACGAG GAGGTGGCAC

+2 S V A H D G A G K R V Y Y L T R D P T T P L A R A A W
6641 TCAGTCGCC ACACGGCC TGGAAAGAG GTCTACTACC TCACCCGTGA CCCTACAACC CCCCTCGCGA GAGCTGCGTG
AGTCAGCGG TGCTGCCGG ACCTTCTCC CAGATGATGG AGTGGCACT GGGATGTTGG GGGGAGCGCT CTCGACGGCAC

+2 E T A R H T P V N S W L G N I I M F A P T L W A R M
6721 GGAGACAGCA AGACACACTC CAGTCAATTC CTGGCTAGGC AACATAATCA TGTTCGCCC CACACTGTGG GCGAGGATGA
CCTCTGCGT TCTGTGTGAG GTCAGTTAAG GACCGATCCG TTGTATTAGT ACAAACGGGG GTGTGACACC CGCTCCTACT

+2 I L M T H F F S V L I A R D Q L E Q A L D C E I Y G A
6801 TACTGATGAC CCATTCTTT AGCGTCCTTA TAGCCAGGGA CCAGCTTGAA CAGGCCCTCG ATTGCGAGAT CTACGGGGCC
ATGACTACTG GGTAAGAAA TCGCAGGAAT ATCGGTCCCT GGTCGAACTT GTCCGGGAGC TAACGCTCTA GATGCCCGG

+2 C Y S I E P L D L P P I I Q R L H G L S A F S L H S Y
6881 TGCTACTCCA TAGAACCCT GGAATCACTT CCAATCACTT AAAGACTCCA TGGCTCAGC GCATTTTCAC TCCACAGTTA
ACGATGAGT ATCTTGGTGA CCTAGATGGA GGTAGTAAG TTTCTCAGT ACCGAGTGC CGTAAAGTG AGGTCTCAAT

FIG. 5L



pCMV-delNS35

+2 S P G E I N R V A A C L R K L G V P P L R A W R H R
6961 CTCTCCAGGT GAAATCAATA GGTGGCCGC ATGCTCAGA AAACCTGGGG TACCGCCCTT GCGAGCTTGG AGACACCGGG
GAGAGGTCCA CTTAGTTAT CCCACCGGG TACGGAGTCT TTTGAACCC ATGGCGGAA CGCTCGAACC TCTGTGGCCC

+2 A R S V R A R L L A R G G R A A I C G K Y L F N W A V
7041 CCCGAGCGT CCGCGTAGG CTTCTGCCA GAGGAGCAG GGTGCCATA TGTGGCAAGT ACCTCTTCAA CTGGCAGTA
GGGCTCGCA GCGCGATCC GAAGACGGT CTCTCCGTC CCGACGGTAT ACACGGTTCA TGGAGAAGTT GACCCGTCAT

+2 R T K L K L T P I A A A G Q L D L S G W F T A G Y S G
7121 AGAACAAAGC TCAAACTCAC TCCAATAGCG GCGCTGGCC AGCTGGACTT GTCCGGCTGG TTCACGGCTG GCTACAGCGG
TCTTCTTTCG AGTTGAGTG AGTTATCGC CGGCGACCGG TCGACCTGAA CAGGCCGACC AAGTCCGAC CGATGTGGCC

+2 G D I Y H S V S H A R P R W I W F C L L L L A A G V
7201 GGGAGACATT TATCACAGCG TGTCTCATGC CCGCCCCCG TGGATCTGTT TTTGCCCTACT CCTGCTTGGT GCAGGGGTAG
CCCTCTGTAA ATAGTGTGC ACAGAGTACG GGCCGGGGG ACCTAGACCA AACGGATGA GGACGAACGA CGTCCCCCAT

+2 G I Y L L P N R
7281 GCATCTACCT CCTCCCCAAC CGATGAAGGT TGGGGTAAAC ACTCCGGCCT AAAAAAATA AAAATCTAG AAAGGCGCGC
CGTAGATGGA GGAGGGGTG GCTACTTCCA ACCCATTTG TGAGGGCGGA TTTTITTTT TTTTITAGATC TTTCCGCGCG

BamHI MluI

7361 CAAGATATCA AGGATCCACT ACGGCTTAGA GCTCGTGAT CAGCCTCGAC TGTGCCCTTCT AGTTGCCAGC CATCTGTGT
GTTCTATAGT TCCTAGGTGA TGCGCAATCT CGAGCGACTA GTCGGAGCTG ACACGGAAGA TCAACGGTCTG GTAGACAACA

7441 TTGCCCTCC CCCGTGCCCT CTTTGACCCT GGAAGGTGCC ACTCCCACTG TCCTTTCCTA ATAAATGAG GAAATTCAT
AACGGGAGG GGGCAGGAA GGAAGTGGCA CCTTCCACGG TGAGGTGAC AGGAAAGGAT TATTTTACTC CTTTAAAGTA

FIG. 5M



pCMV-delNS35

7521 CGCATTGTCT GAGTAGGTGT CATTCTATTG TGGGGGGTGG GGTGGGGCAG GACAGCAAGG GGGAGGATTG GGAAGACAAT
GCGTAAACAGA CTCATCCACA GTAAGATAAG ACCCCCCACC CGACCCCGTC CTGTCGTTCC CCGTCCCTAAC CCTTCTGTTA

7601 AGCAGGCATG CTGGGAGCT CTTCCGCTTC CTGCGTCACT GACTCGCTGC GCTCGGTGCT GCTCGGTGCGG CGAGCGGTAT
TCGTCCGTAC GACCCCTCGA GAAGGCGAAG GAGCGAGTGA CTGAGCGACG CGAGCCAGCA AGCCGACGCC GCTCGCCATA

7681 CAGTCACTC AAAGGCGGTA ATACGGTTAT CCACAGATC AGGGGATAAC GCAGGAAAGA ACATGTGAGC AAAAGGCCAG
GTCGAGTGAG TTTCCGCCAT TATGCCAATA GGTGCTTTAG TCCCCTATTG CGTCCCTTCT TGTACACTCG TTTTCCGGTC

7761 CAAAAGGCCA GGAACCGTAA AAAGGCCGGC TTGCTGGGT TTTTCCATAG GCTCCGCCCC CCGTACGAGC ATCACAATAA
GTTTTCCGGT CCTTGGCATT TTTCCGCCGC AACGACCGCA AAAAGGTATC CGAGCGGGGG GGACTGCTCG TAGTGTTTTT

7841 TCGACGCTCA AGTCAGAGGT GCGGAAACCC GACAGGACTA TAAAGATACC AGCGTTTCC CCTCGGAAGC TCCCTCGTGC
AGCTGCGAGT TCAGTCTCCA CCGCTTTGGG CTGTCTTGAT ATTCTATGG TCCGCAAGG GGCACCTTCG AGGAGCACG

7921 GCTCTCCTGT TCCGACCCTG CCGCTTACCG GATACCTGTC CGCCTTTCTC CCTTCGGGAA GCGTGGCGCT TTCTCAATGC
CGAGAGGACA AGGCTGGGAC GCGGAATGGC CTATGGACAG GCGGAAAGAG GGAAGCCCTT CGCACCCGGA AAGAGTTACG

8001 TCACGCTGTA GGTATCTCAG TTCGGTGTAG GTCGTTGCT CCAAGCTGGG CTGTGTGCAC GAACCCCCCG TTCAGCCCCGA
AGTCCGACAT CCATAGAGTC AAGCCACATC CAGCAAGCGA GGTTCGACCC GACACACGTG CTGGGGGGC AAGTCGGGCT

8081 CCGCTGGCC TTATCCGGTA ACTATCGTCT TCAGTCCAAC CCGTAAGAC ACGACTTATC GCCACTGGCA GCAGCCACTG
GGGACGCCG AATAGGCCAT TGATAGCAGA ACTCAGGTTG GGCCATTCTG TGCTGAATAG CCGTGACCGT CGTCGGTGAC

8161 GTAACAGGAT TAGCAGAGCG AGGTATGTAG GCGGTGCTAC AGAGTTCTTG AAGTGTGCG CTAACACTAGG CTACACTAGA
CATTGTCCTA ATCGTCTCGC TCCATACATC CGCCACGATG TCTCAAGAAC TTCACCACCG GATTGATGCC GATGTGATCT

FIG. 5N



pCMV-delNS35

8241	AGGACAGTAT TTGGTATCTG CGCTCTGCTG AAGCCAGTTA CTTTCGGAAA AAGAGTTGGT AGCTCTTGAT CCGGCAAAACA TCCTGTGATA AACCATAGAC GCGAGACGAC TTCGGTCAAT GGAAGCCTTT TTCTCAACCA TCGAGAACTA GGCCGTTTGT
8321	AACCACCGCT GGTAGCGGTG GTTTTTTTGT TTGCAAGCAG CAGATTACGC GCAGAAAAAA AGCATCTCAA GAAGATCCTT TTGGTGGCGA CCATCGCCAC CAAAAAACA AACGTGCTC GTCTAATGCG CGTCTTTTTT TCCTAGAGTT CTTCTAGGAA
8401	TGATCTTTTC TACGGGGTCT GACGCTCAGT GGAACGAAAA CTCACGTTAA GGGATTTTGG TCATGAGATT ATCAAAAAGG ACTAGAAAAG ATGCCCCAGA CTGCGAGTCA CCTTGCTTTT GAGTGCAATT CCTAAAAACC AGTACTCTAA TAGTTTTTCC
8481	ATCTTCACCT AGATCCTTTT AAATTAAAA TGAAGTTTAA AATCAATCTA AAGTATATAT GAGTAAACTT GGTCTGACAG TAGAAGTGA TCTAGGAAA TTAAATTTTT ACTTCAAAAT TTAGTTAGAT TTCAATATA CTCATTTGAA CCAGACTGTC
8561	TTACCAATGC TTAATCAGTG AGGCACCTAT CTCAGCGATC TGCTATTTC GTTCATCCAT ACTTGCCTGA CTCCTCCGTG AATGGTTACG AATTAGTCAC TCGGTGGATA GAGTCGTAG ACAGATAAAG CAAGTAGGTA TCAACGGACT GAGGGGCAGC
8641	TGTAGATAAC TACGATACGG GAGGGCTTAC CATCTGGCCC CAGTGCTGCA ATGATACCGC GAGACCCACG CTCACCGGCT ACATCTATTG ATGCTATGCC CTCCGGAATG GTACACGGGG GTCTATGGCG CTCTGGGTGC GAGTGGCCGA
8721	CCAGATTTAT CAGCAATAAA CCAGCCAGCC GGAAGGGCCG AGCGCAGAAG TGGTCTGCA ACTTTATCCG CCTCCATCCA GGTCTAAATA GTCGTTATT GTTCGGTGGG CTTCCCGGC TCGGTCTTC ACCAGGACGT TCAATAGGC GGAGGTAGGT
8801	GTCTATTAAAT TGTTCGGGG AAGCTAGAGT AAGTAGTTCC CCAGTTAATA GTTTGGCGCA CGTTGTTGCC ATTGCTACAG CAGATAATTA ACAACGGCCC TTCGATCTCA TTCAATCAAG GGTCAATTAT CAAACGGCTT GCAACAACGG TAACGATGTC
8881	GCATCGTGGT GTCACGCTCG TCGTTTGGTA TGGTTTCATT CAGCTCCGGT TCCCAACGAT CAAGGGCAGT TACATGATCC CGTAGCACCA CAGTGGGAGC AGCAACCAT ACCGAAGTAA GTCGAGGCCA AGGTTTGCTA GTTCGGCTCA ATGTACTAGG

FIG. 50



pCMV-delINS35

8961 CCCATGTTGT GCAAAAAGC GGTAGCTCC TTGGTCTCTC CGATCGTTGT CAGAAGTAAG TTGGCCGCAG TGTATCACT
GGGTACAACA CGTTTTTTCG CCAATCGAGG AAGCCAGGAG GCTAGCAACA GTCTTCATTC AACGGCGTC ACAATAGTGA

9041 CATGGTTATG GCAGCACTGC ATAATTCTCT TACTGTCATG CCATCCGTAA GATGCTTTTC TGTGACTGGT GAGTACTCAA
GTACCAATAC CGTCGTGACG TATTAAGAGA ATGACAGTAG GGTAGGCATT CTACGAAAAG ACACTGACCA CTCATGAGTT

9121 CCAAGTCATT CTGAGAAATAG TGTATGCGGC GACCGAGTTG CTCTTGCCCG GCGTCAATAC GGGATAATAC CGGGCCACAT
GGTTCAGTAA GACTCTTATC ACATACGCCG CTGGCTCAAG GAGAACGGGC CGCAGTTATG CCTATTATG GCGGGGTGA

9201 AGCAGAACTT TAAAAGTGCT CATCATTTGA AAACGTTCTT CGGGGCGGAA ACTCTCAAGG ATCTTACCGC TGTGAGATC
TCGTCITTGAA ATTTTCACGA GTAGTAACCT TTGCAAGAA GCCCGGCTTT TGAGAGTTCC TAGAATGGCG ACAACTCTAG

9281 CAGTTCGATG TAACCCACTC GTGCACCCAA CTGATCTTCA GCATCTTTTA CTTTCACCCAG CGTTTCTGGG TGAGCAAAAA
GTCAAGCTAC ATTGGGTGAG CACGTGGGTT GACTAGAAGT CGTAGAAAT GAAAGTGCTC GCAAAGACCC ACTCGTTTTT

9361 CAGGAAGCA AAATGCCGCA AAAAAGGGA TAAGGGCGAC ACGGAAATGT TGAATACTCA TACTCTTCCT TTTTCAATAT
GTCCTTCCGT TTACGGCGT TTTTTCGCTT ATTCCGCTG TGCCTTTACA ACTTATGAGT ATGAGAAGGA AAAAGTTATA

9441 TATTGAAGCA TTTATCAGG TTATTGTCTC ATGAGCGGAT ACATATTTGA ATGTATTTAG AAAAATAAAC AAATAGGGGT
ATAACTTCGT AAATAGTCCC AATAACAGAG TACTCGCCTA TGTATAAACT TACATAAATC TTTTATTTG TTTATCCCCA

9521 TCCGGGACA TTTCCCCGAA AAGTGCACC TGACGTCTAA GAAACCATTA TTAATCATGAC ATTAACCTAT AAAAATAGGC
AGCGCGTGT AAAGGGCTT TTCACGGTGG ACTGCAGATT CTTTGGTAAT AATAGTACTG TAATTGGATA TTTTATCCG

9601 GTATCAGGAG GCCCTTTCTG C
CATAGTCTC CGGAAAGCA G

FIG. 5P

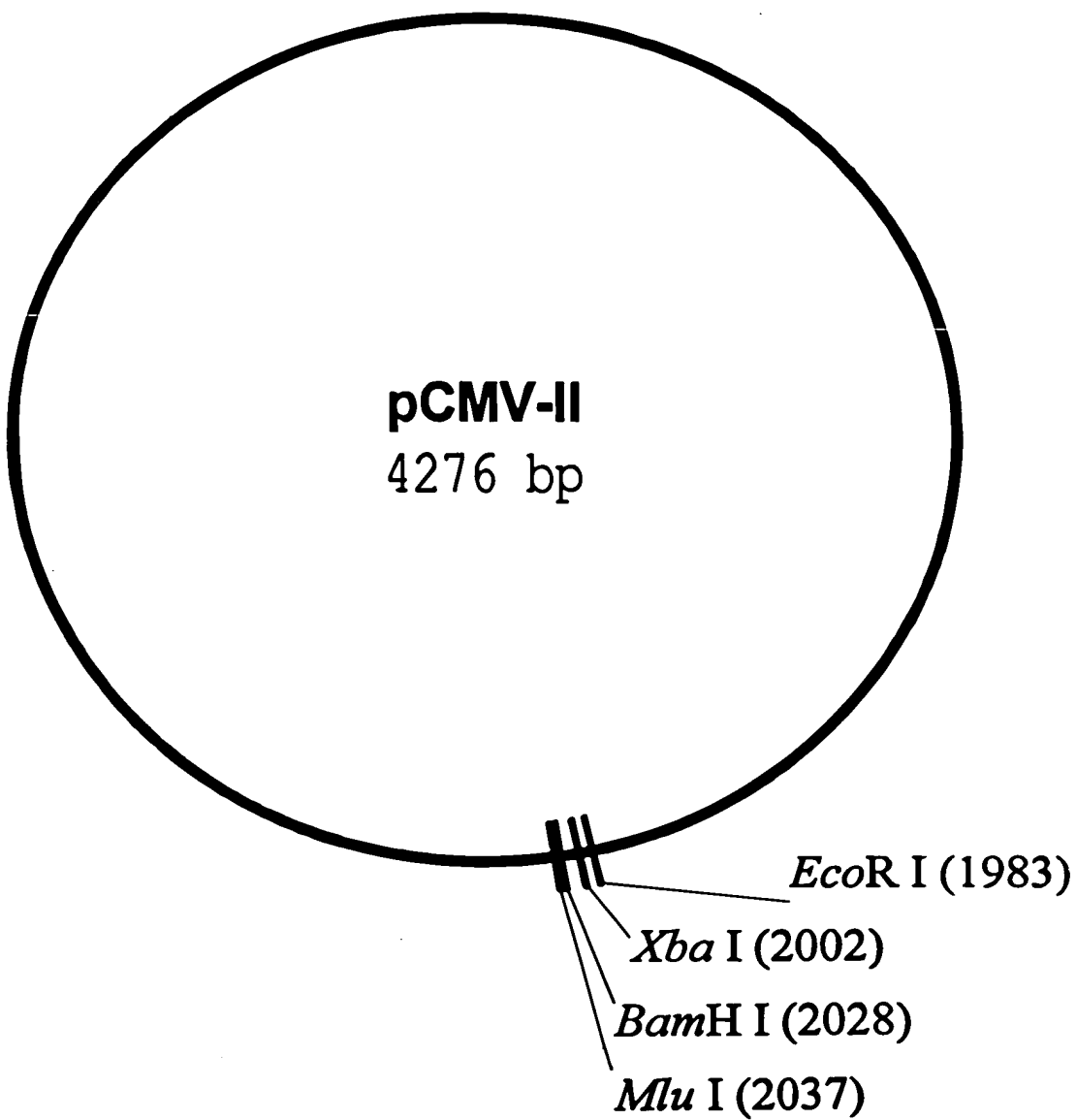


FIG. 6



pCMV-II

1 TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG GAGACGGTCA CAGCTTGTCT GTAAGCGGAT
AGCGCGCAA GCCACTACTG CCACTTTTGG AGACTGTGA CGTGAGGCG CTCTGCCAGT GTCGAACAGA CATTCGCCTA

81 GCCGGGAGCA GACAAGCCCG TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG TCGGGGGCTGG CTTAACATATG CGGCATCAGA
CGGCCCTCGT CTGTTCCGGC AGTCCCGCGC AGTCCCGCGC AACCGCCAC AGCCCGGACC GAATTGATAC GCCGTAGTCT

161 GCAGATTGTA CTGAGAGTGC ACCATATGAA GCTTTTGGCA AAAGCTAGG CCTCCAAAA AGCCTCCTCA CTACTTCTGG
CGTCTAACAT GACTCTCAG TGGTATACTT CGAAAAACGT TTTCCGATCC GGAGGTTTT TCGGAGGAGT GATGAAGACC

241 AATAGCTCAG AGCCCGAGGC GGCCTCGGC TCTGCATAA TAAAAAAAT TAGTCAGCCA TGGGGCGGAG AATGGCGGGA
TTATCGAGT TCCGGCTCCG CCGGAGCCGG AGACGTATTT ATTTTTTTTA ATCAGTCGGT ACCCGGCTC TTACCCGCT

321 ACTGGCGGG GAGGGAATTA TTGGCTATTG GCCATTGCAT ACCTGTATC TATATCATAA TATGTACATT TATATTGGCT
TGACCCGCCC CTCCTTAAT AACCGATAAC CCGTAACGTA TCGAACATAG ATATAGTATT ATACATGTAA ATATAACCGA

401 CATGTCCAAT ATGACCGCCA TGTTGACATT GATTATTGAC TAGTTATTAA TAGTAATCAA TTACGGGGTC ATTAGTTCAT
GTACAGGTTA TACTGGCGGT ACAACTGTAA CTAATAACTG ATCAATAATT ATCAATTAGT AATGCCCCAG TAATCAAGTA

481 AGCCCATATA TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC CCGCCCCATT
TCGGGTATAT ACCTCAAGC GCAATGTATT GAATGCCATT TACCGGGCGG ACCGACTGGC GGGTTGCTGG GGGCGGGTAA

561 GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTACGGT
CTGCAGTTAT TACTGCATAC AAGGTATCA TTGCGGTTAT CCTGAAAGG TAACTGCAGT TACCCACCTC ATAAATGCCA

641 AAAC TGCCCA CATCAAGTGT ATCATATGCC AACTCCGCCC CCTATTGACG TCAATGACGG TAAATGGCCC
TTTGACGGGT GAACCGTCAT GTAGTTCACA TAGTATACGG TTCAGGCGGG GGATAACTGC AGTTACTGCC ATTTACCGGG

FIG. 7A



pcMV-II

721	GCCTGGCATT ATGCCCAGTA CATGACCTTA CGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA TCGCTATTAC CGGACCGTAA TACGGGTCAAT GACTGGAAT GCCCTGAAAG GATGAACCGT CATGTAGATG CATAATCAGT AGCGATAATG
801	CATGGTGATG CGGTTTTGGC AGTACACCAA TGGGGGTGGA TAGGGTTTG ACTCAGGGG ATTTCCAAAGT CTCCACCCCA GTACCACTAC GCCAAAACCG TCATGTGGT ACCCGCACCT ATCGGCAAC TGAGTGCCCC TAAAGTTCA GAGGTGGGT
881	TTGACGTCAA TGGGAGTTTG TTTTGGCACC AAAATCAACG GGACTTTCCA AAATGTCGA ATAACCCCGC CCCGTTGACG AACTGCAGT ACCCTCAAAC AAAACCGTGG TTTTAGTTGC CCTGAAAGGT TTTACAGCAT TATTGGGGCG GGGCAACTGC
961	CAAATGGCG GTAGGCGTGT ACGGTGGGAG GTCTATATAA GCAGAGCTCG TTTAGTGAAC CGTCAGATCG CCTGGAGACG GTTTACCCGC CATCCGCACA TGCCACCCTC CAGATATATT CGTCTGAGC AAATCACTTG GCAGTCTAGC GGACCTCTGC
1041	CCATCCACGC TGTTTTGACC TCCATAGAAG ACACCGGGAC CGATCCAGCC TCCGGGGCGG GGAACGGTGC ATTGGAACGC GGTAGGTGCG ACAAACCTGG AGGTATCTTC TGTGGCCCTG GCTAGGTCGG AGCGCCCGC CCTTGCCACG TAACCTTGGC
1121	GGATTCCCG TGCCAAGAGT GACGTAAGTA CCGCCTATAG ACTCTATAGG CACACCCCTT TGGCTCTTAT GCATGCTATA CCTAAGGGCG ACGGTTCTCA CTGCATTCA CTGGGATATC TGAGATATCC GTGTGGGGA ACCGAGAATA CGTACGATAT
1201	CTGTTTTTGG CTGGGGCCT ATACACCCCG GCTCCTTATG CTATAGGTGA TGGTATAGCT TAGCCTATAG GTGTGGGTTA GACAAAAACC GAACCCCGGA TATGTGGGG CGAGGAATAC GATATCCACT ACCATATCGA ATCGGATATC CACACCCAAT
1281	TTGACCATT TTAGCCACTC CCTATTGGT GACGATACTT TCCATTACTA ATCCATAACA TGGCTCTTTG CCACAACAT AACTGGTAAT AACTGGTGAG GGGATAACCA CTGCTATGAA AGGTAATGAT TAGGTATTGT ACCGAGAAAC GGTGTTGATA
1361	CTCTATTGGC TATATGCCAA TACTCTGTCC TTCAGAGACT GACACGGACT CTGTATTTT ACAGATGGG GTCCATTAT GAGATAACCG ATATACGGTT ATGAGACAGG AAGTCTCTGA CTGTGCCTGA GACATAAAAA TGTCTACCC CAGGTAAATA

FIG. 7B



pCMV-II

```
1441 TATTACAAA TTCACATATA CAACAAGCC GTCCCGCGTG CCGCAGTTT TTATTAAACA TAGCGTGGGA TCTCCGACAT
    AATAATGTTT AACTGTATAT GTTGTGCGG CAGGGGGCAC GGGCGTCAAA AATAATTGT ATCGCACCT AGAGGCTGTA

1521 CTCGGGTACG TGTTCGGAC ATGGGCTCTT CTCCGTAGC GCGGAGCTT CCACATCCGA GCCCTGGTCC CATCCGTCCA
    GAGCCCATGC ACAAGGCCGTG TACCCGAGAA GAGGCCATCG CCGCTCGAA GGTGTAGGCT CGGACCCAGG GTAGGCAGGT

1601 GCGGCTCATG GTCGCTCGG AGTCCTTGC TCCTAACAGT GGAGGCCAGA CTTAGGCACA GCACAATGCC CACCACCACC
    CGCCGAGTAC CAGCGAGCCG TCGAGGAACG AGGATTGTC AATTCTCGA CCTCCGGTCT GAATCCGTGT CTTCTTACGG GTGGTGGTGG

1681 AGTGTGCCG ACAAGGCCGT GCGGTAGGG TATGTGTCTG AAAATGAGCT CGGAGATTGG GCTCGCACCT GGACGCAGAT
    TCACACGGCG GTTTCGGCA CCGCCATCCC ATACACAGAC TTTTACTCGA GCCTCTAAC CCAGCGTGA CCTGCCGTCTA

1761 GGAAGACTTA AGGCAGCGG AGAAGAAGT GCAGGCAGCT GAGTTGTTGT ATTCTGATAA GAGTCAAGGG TAACCTCCCT
    CCTTCTGAAT TCCGTCGCCG TCTTCTTCTA CTTCTGCGA CTCAACAACA TAAGACTATT CTCAGTCTCC ATTGAGGGCA

1841 TGCGGTGCTG TTAACGGTGG AGGCAGTGT AGTCTGAGCA GTACTCGTTG CTGCCGGCGG CGCCACCAGA CATAATAGCT
    ACGCCACGAC AATTGCCACC TCCCGTCACA TCAGACTCGT CATGAGCAAC GACGGCGCGC CCGGTGGTCT GTATTATCGA

1921 GACAGACTAA CAGACTGTTT CTTTCCATGG GTCTTTTCTG CAGTCACCGT CTTGACCTA AGAATTCAGA CTCGAGCAAG
    CTGCTGATT GTCTGACAAG GAAAGGTAAC CAGAAAAGAC GTCAGTGGCA GCAGCTGGAT TCTTAAGTCT GAGCTCGTTC

2001 TCTAGAAAGG GCGGCCAAGA TATCAAGGAT CCACTACGCG TTAGAGCTCG CTGATCAGCC TCGACTGTGC CTTCTAGTTG
    AGATCTTCC GCGCGGTTCT ATAGTTCTTA GGTGATGCC AATCTCGAGC GACTAGTCGG AGCTGACAG GAAGATCAAC
```

XbaI BamHI MluI EcoRI

FIG. 7C



pCMV-II

2081	CCAGCCATCT GTTCTTTGCC CCTCCCCCGT GCCTTCCTTG ACCCTGGAAG GTGCCACTCC CACTGTCTTT TCCTAATAAAA GGTCGGTAGA CAACAAACGG GGAGGGGCA CGGAAGGAAC TGGGACCTTC CACGGTGAGG GTGACAGGAA AGGATTATTT
2161	ATGAGGAAAT TGCATCGCAT TGTCTGAGTA GGTGTCATTG TATTCTGGG GGTGGGGTGG GGCAGGACAG CAAGGGGGAG TACTCCCTTA ACGTAGCGTA ACAGACTCAT CCACAGTAAG ATAAGACCCC CCACCCCACC CCGTCTCTGC GTTCCCCCTC
2241	GATTGGGAAG ACAATAGCAG GCATGCTGGG GAGCTCTTCC GCTTCTCTGC TCACTGACTC GCTGGGCTCG GTCGTTCCGC CTAACCCCTC TGTATCGTC CGTACGACCC CTCGAGAAG CGAAGGAGCG AGTGAAGTGC CGACGGGAGC CAGCAAGCCG
2321	TGCGGGGAGC GGTATCAGCT CACTCAAAG CGCTAATACG GTTATCCACA GAATCAGGG ATAACGCAGG AAAGAACATG ACGCCGCTCG CCATAGTCGA GTGAGTTTCC GCCATTATGC CAATAGGTGT CTTAGTCCCC TATTGGGTCC TTTCTTTGATC
2401	TGAGCAAAAG GCCAGCAAAA GGCCAGGAAC CCTAAAAAG CCGGTTGCT GCGCTTTTTC CATAGGCTCC GCCCCCCTGA ACTCGTTTC CCGTCTGTTT CCGGTCTCTG GCATTTTTC GCGCAACGA CCGCAAAAAG GTATCCGAGG CGGGGGGACT
2481	CGAGCATCAC AAAAATCGAC GCTCAAGTCA GAGGTGGCGA AACCCGACAG GACTATAAG ATACCAGGCG TTTCCCGCTG GCTCGTAGTG TTTTAGCTG CGAGTTTCACT CTCCACCGCT TTGGGCTGTC CTGATATTTC TATGGTCCGC AAAGGGGGAC
2561	GAAGCTCCCT CGTGGGCTCT CTGTCTCGGA CCTGCGCGCT TACCGGATAC CTGTCCGCTT TCTCCCTTC GGGAAGCGTG CTTCGAGGGA GCACGGGAGA GAACAAGGCT GGCACGGCGA ATGGCTATG GACAGGGCGA AAGAGGGAAG CCTTTCGCAC
2641	GGGCTTTCTC AATGCTCAG CTGTAGGTAT CTCAGTTCCG TGTAGTCTG TCGTCCAAG CTGGGCTGTG TGCACGAACC CGCGAAAGAG TTAGGAGTGC GACATCCATA GAGTCAAGCC ACATCCAGCA AGCGAGCTTC GACCCGACAC ACGTGTCTTG
2721	CCCCGTTTCA CCCGACCGCT GCGCCTTATC CCGTAACTAT CGTCTTGAGT CCAACCCCGT AAGACACGAC TTATCGCCAC GGGGCAAGTC GGGCTGGCGA CGCGGAATAG GCCATTGATA GCAGAACTCA GCTTGGGCA TCTGTGCTG AATAGCGGTG

FIG. 7D



pCMV-II

2801 TGGCAGCAGC CACTGGTAAC AGGATTAGCA GAGCGAGGTA TGTAGGCGGT GCTACAGAGT TCTTGAAGTG GTGGCCTAAC
ACCGTCGTGG GTGACCAATG TCCTAATCGT CTCGCTCCAT ACATCCGCCA CGATGTCTCA AGAATCTCAC CACCGGATTG

2881 TAGGGCTACA CTAGAAGGAC AGTATTGGT ATCTGGCTC TGCTGAAGCC AGTTACCTTC GGAAAAAGAG TTGCTAGCTC
ATGCCGATGT GATCTTCTG TCATAAACCA TAGACGGGAG ACGACTGGG TCAATGGAAG CCTTTTCTC AACCATCGAG

2961 TTGATCCGGC AAACAAACCA CCGCTGGTAG CCGTGGTTTT TTTGTTTGCA AGCAGCAGAT TACGGGCAGA AAAAAAGGAT
AACTAGGCCG TTTGTTTGGT GCGACCATC GCCACCAAAA AAACAAAGCT TCGTGTCTA ATGCGGCTCT TTTTTCCTA

3041 CTCAAGAAGA TCCTTTGATC TTTTCTACGG GGTCTGACGC TCAGTGAAC GAAAACTCAC GTTAAAGGGAT TTTGCTCATG
GAGTTCTTCT AGGAAACTAG AAAAGATGCC CCAGACTGG AGTCACCTG CTTTIGAGTG CAATTCCCTA AAACAGTAG

3121 AGATTATCAA AAAGGATCTT CACCTAGATC CTTTTAAAT AAAATGAAG TTTTAAATCA ATCTAAAGTA TATATGAGTA
TCTAATAGTT TTTTCTAGAA GTGGATCTAG GAAAATTAA TTTTACTTC AAAATTAGT TAGATTTTCT ATATACTCAT

3201 AACTTGGTCT GACAGTTACC AATGCTTAAT CAGTGAGGCA CCTATCTCAG CGATCTGTCT ATTTCTGTTCA TCCATAGTTG
TTGAACCAGA CTGTCAATGG TTACGAATTA GTCACTCCGT GGATAGAGTC GCTAGACAGA TAAAGCAAGT AGCTATCAAC

3281 CCTGACTCCC CGTCGTGTAG ATAACTACGA TACGGGAGGG CTTACCATCT GCGCCCAAGTG CTGCAATGAT ACCGGGAGAC
GGACTGAGGG GCAGCACATC TATTGATGCT TG CCTCCC GAATGTAGA CCGGGGTAC GACGTTACTA TGGCGCTCTG

3361 CCACGCTCAC CGGCTCCAGA TTTATCAGCA ATAAACAGC CAGCGGGAAG GCGCGAGCGG AGAAGTGGTC CTGCAACTTT
GGTCCGAGTG GCGGAGGTCT AAATAGTCTG TATTGGTGG GTGCGGCTTC CCGGCTCGG TCTTCACCAG GACGTTGAAA

3441 ATCCGCCCTCC ATCCAGTCTA TTAATTGTTG CCGGGAAGCT AGAGTAAGTA GTTCGCCAGT TAATAGTTTG CGCAACGTTG
TAGGCGGAGG TAGGTCAGAT AATTAACAAC GGCCTTCCA TCTCATTCAT CAAGCGGTCA ATTATCAAC GCGTTGCAAC

FIG. 7E



pCMV-II

3521 TTGCCATTGC TACAGGCATC GTGGTGTAC GCTCGTCGTT TGGTATGGCT TCATTACAGCT CCGGTTCCCA ACGATCAAGG
AACGGTAACG ATGTCCGTAG CACCACAGTG CGACAGCAA ACCATACCGA AGTAAGTCGA GGCCAAGGCT TGCTAGTTCC

3601 CGAGTTACAT GATCCCCCAT GTTGIGCAAA AAAGGGTTA GCTCCTTCGG TCCTCCGATC GTTGTACAGAA GTAAGTTGGC
GGTCAATGTA CTAGGGGGTA CAACACGTTT TTTCCGCAAT CGAGGAAGCC AGGAGGCTAG CAACAGTCTT CATTCAACCG

3681 CGCAGTGTTA TCACTCATGG TTATGGCAGC ACTGCATAAT TCTCTTACTG TCATGCCATC CGTAAGATGC TTTTCTGTGA
GGGTACAAT AGTGAGTACC AATACCGTCG TGACGTATTA AGAGAAATGAC AGTACGGTAG GCATTCTACG AAAAGACACT

3761 CTGGTGACTA CTCAACCAAG TCATTCTGAG AATAGTGTAT GCGGCGACCG AGTTGCTCTT GCCCGGCGTC AATACGGGAT
GACCACTCAT GAGTTGGTTC AGTAAGACTC TTATCACATA CGCCGCTGGC TCAACCGAGAA CCGGCCGCAG TTATGCCCTA

3841 AATACCGCGC CACATAGCAG AACTTTAAAA GTGCTCATCA TTGGAACACG TTCTTCGGGG CGAAAACTCT CAAGGATCTT
TTATGGCGCG GTGTATCGTC TTGAAATTTT CACGAGTAGT AACCTTTTGC AAGAAGCCCC GCTTTTGAGA GTTCCTAGAA

3921 ACCGCTGTTG AGATCCAGTT CGATGTAACC CACTCGTGCA CCCAACTGAT CTTCAGCATC TTTTACTTTC ACCAGCGTTT
TGGCGACAAC TCTAGGTCAA GCTACATTGG GTGAGCACGT GGTTGACTA GAAGTCGTAG AAAATGAAAG TGGTCGCAAA

4001 CTGGTGAGC AAAACACGGA AGGCAAAATG CCGCAAAAAA GGAATAAGG GCGACACGGA AATGTTGAAT ACTCATACTC
GACCCACTCG TTTTGTCTT TCCGTTTAC GCGGTTTTTT CCCTTATTCC CGCTGCGCT TTACAACTTA TGAGTATGAG

4081 TTCCTTTTTC AATATTATTG AAGCATTTAT CAGGGTTATT GTCTCATGAG CGGATACATA TTTGAATGTA TTTAGAAAAA
AAGGAAAAAG TTATAATAAC TTCGTAATA GTCCCAATAA CAGAGTACTC GCCTATGTAT AAACCTTACAT AAATCTTTTT

4161 TAAACAAATA GGGTTCCGC GCACATTTCC CCGAAAAGTG CCACCTGAGC TCTAAGAAAC CATTATTATC ATGACATTAA
ATTTGTTTAT CCCCAAGGCG CGTGTAAGG GGCTTTTTCAC GGTGGAAGTGC AGATTCTTTG GTAATAATAG TACTGTAATT

4241 CCTATAAAAA TAGGCGTATC ACGAGGCCCT TTCGTC
GGATATTTT ATCCGCATAG TGCTCCGGGA AAGCAG

FIG. 7F

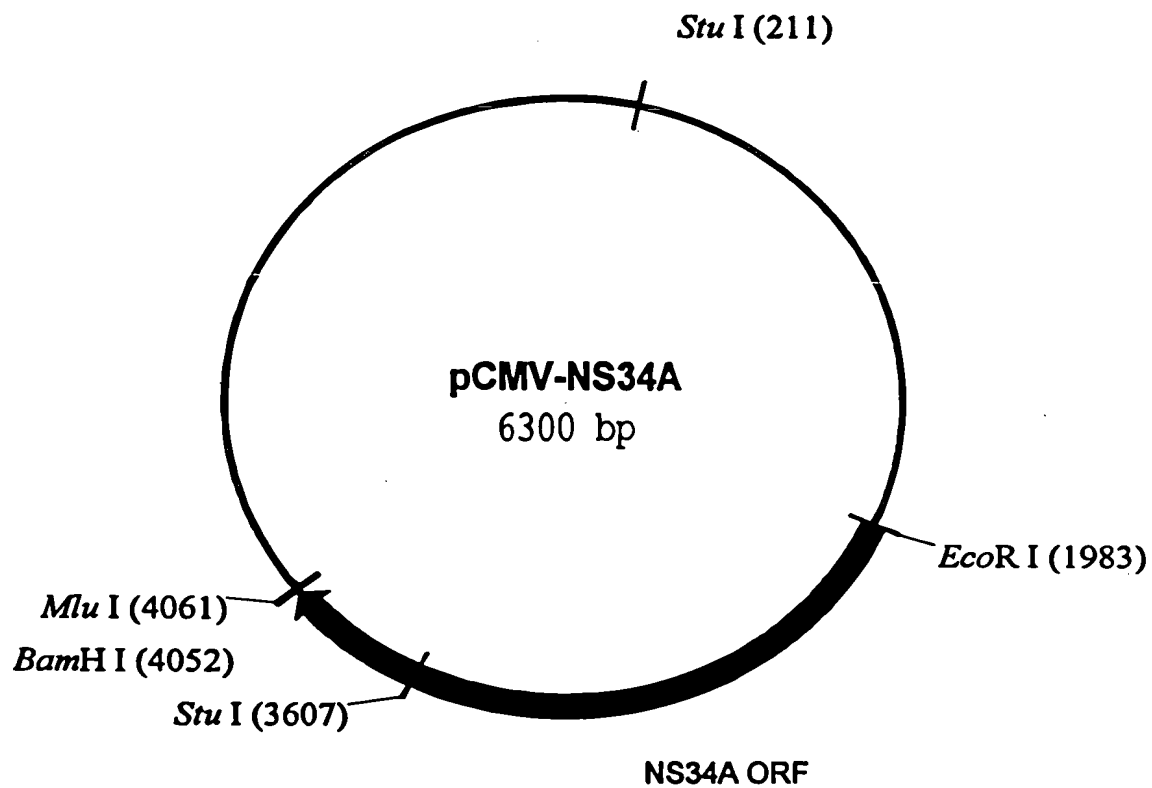
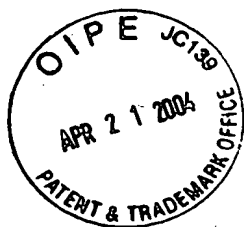


FIG. 8



pCMV-NS34A

1	TCGCGCGTTT	CGGTGATGAC	GGTGAAAACC	TCTGACACAT	GCAGCTCCCG
	AGCGCGCAAA	GCCACTACTG	CCACTTTTGG	AGACTGTGTA	CGTCGAGGGC
51	GAGACGGTCA	CAGCTTGTCT	GTAAGCGGAT	GCCGGGAGCA	GACAAGCCCG
	CTCTGCCAGT	GTCGAACAGA	CATTGCGCTA	CGGCCCTCGT	CTGTTCTGGG
101	TCAGGGCGCG	TCAGCGGGTG	TTGGCGGGTG	TCGGGGCTGG	CTTAACATATG
	AGTCCCGCGC	AGTCGCCCAC	AACCGCCCAC	AGCCCCGACC	GAATTGATAC
151	CGGCATCAGA	GCAGATTGTA	CTGAGAGTGC	ACCATATGAA	GCTTTTTTGA
	GCCGTAGTCT	CGTCTAACAT	GACTCTCACG	TGGTATACTT	CGAAAAACGT
StuI					
~~~~~					
201	AAAGCCTAGG	CCTCCAAAAA	AGCCTCCTCA	CTACTTCTGG	AATAGCTCAG
	TTTCGGATCC	GGAGGTTTTT	TCGGAGGAGT	GATGAAGACC	TTATCGAGTC
251	AGGCCGAGGC	GGCCTCGGCC	TCTGCATAAA	TAAAAAAAT	TAGTCAGCCA
	TCCGGCTCCG	CCGGAGCCGG	AGACGTATTT	ATTTTTTTTA	ATCAGTCGGT
301	TGGGGCGGAG	AATGGGCGGA	ACTGGGCGGG	GAGGGAATTA	TTGGCTATTG
	ACCCCGCCTC	TTACCCGCCT	TGACCCGCCC	CTCCCTTAAT	AACCGATAAC
351	GCCATTGCAT	ACGTTGTATC	TATATCATAA	TATGTACATT	TATATTGGCT
	CGGTAACGTA	TGCAACATAG	ATATAGTATT	ATACATGTAA	ATATAACCGA
401	CATGTCCAAT	ATGACCGCCA	TGTTGACATT	GATTATTGAC	TAGTTATTAA
	GTACAGGTTA	TACTGGCGGT	ACAACGTAA	CTAATAACTG	ATCAATAATT
451	TAGTAATCAA	TTACGGGGTC	ATTAGTTCAT	AGCCCATATA	TGGAGTTCCG
	ATCATTAGTT	AATGCCCCAG	TAATCAAGTA	TCGGGTATAT	ACCTCAAGGC
501	CGTTACATAA	CTTACGGTAA	ATGGCCCGCC	TGGCTGACCG	CCCAACGACC
	GCAATGTATT	GAATGCCATT	TACCGGGCGG	ACCGACTGGC	GGGTTGCTGG
551	CCCGCCCAT	GACGTCAATA	ATGACGTATG	TTCCTATAGT	AACGCCAATA
	GGGCGGGTAA	CTGCAGTTAT	TACTGCATAC	AAGGGTATCA	TTGCGGTTAT
601	GGGACTTTCC	ATTGACGTCA	ATGGGTGGAG	TATTTACGGT	AACTGCCCCA
	CCCTGAAAGG	TAACTGCAGT	TACCCACCTC	ATAAATGCCA	TTTGACGGGT
651	CTTGGCAGTA	CATCAAGTGT	ATCATATGCC	AAGTCCGCCC	CCTATTGACG
	GAACCGTCAT	GTAGTTCACA	TAGTATACGG	TTCAGGCGGG	GGATAACTGC
701	TCAATGACGG	TAAATGGCCC	GCCTGGCATT	ATGCCCAGTA	CATGACCTTA
	AGTTACTGCC	ATTTACCGGG	CGGACCGTAA	TACGGGTCAT	GTACTGGAAT
751	CGGGACTTTC	CTACTTGGCA	GTACATCTAC	GTATTAGTCA	TCGCTATTAC
	GCCCTGAAAG	GATGAACCGT	CATGTAGATG	CATAATCAGT	AGCGATAATG
801	CATGGTGATG	CGGTTTTTGG	AGTACACCAA	TGGGCGTGGA	TAGCGGTTTG
	GTACCACTAC	GCCAAAACCG	TCATGTGGTT	ACCCGCACCT	ATCGCCAAAC
851	ACTCACGGGG	ATTTCCAAGT	CTCCACCCCA	TTGACGTCAA	TGGGAGTTTG
	TGAGTGCCCC	TAAAGGTTCA	GAGGTGGGGT	AACTGCAGTT	ACCCTCAAAC

FIG. 9A



pCMV-NS34A

901 TTTTGGCACC AAAATCAACG GGACTTTCCA AAATGTCGTA ATAACCCCGC  
AAAACCGTGG TTTTAGTTGC CCTGAAAGGT TTTACAGCAT TATTGGGGCG

951 CCCGTTGACG CAAATGGGCG GTAGGCGTGT ACGGTGGGAG GTCTATATAA  
GGGCAACTGC GTTTACCCGC CATCCGCACA TGCCACCCTC CAGATATATT

1001 GCAGAGCTCG TTTAGTGAAC CGTCAGATCG CCTGGAGACG CCATCCACGC  
CGTCTCGAGC AAATCACTTG GCAGTCTAGC GGACCTCTGC GGTAGGTGCG

1051 TGTTTTGACC TCCATAGAAG ACACCGGGAC CGATCCAGCC TCCGCGGCCG  
ACAAAACCTGG AGGTATCTTC TGTGGCCCTG GCTAGGTCGG AGGCGCCGGC

1101 GGAACGGTGC ATTGGAACGC GGATTCCCCG TGCCAAGAGT GACGTAAGTA  
CCTTGCCACG TAACCTTGCG CCTAAGGGGC ACGGTTCTCA CTGCATTCTAT

1151 CCGCCTATAG ACTCTATAGG CACACCCCTT TGGCTCTTAT GCATGCTATA  
GGCGGATATC TGAGATATCC GTGTGGGGAA ACCGAGAATA CGTACGATAT

1201 CTGTTTTTGG CTTGGGGCCT ATACACCCCC GCTCCTTATG CTATAGGTGA  
GACAAAAACC GAACCCCGGA TATGTGGGGG CGAGGAATAC GATATCCACT

1251 TGGTATAGCT TAGCCTATAG GTGTGGGTTA TTGACCATTA TTGACCACTC  
ACCATATCGA ATCGGATATC CACACCCAAT AACTGGTAAT AACTGGTGAG

1301 CCCTATTGGT GACGATACTT TCCATTACTA ATCCATAACA TGGCTCTTTG  
GGGATAACCA CTGCTATGAA AGGTAATGAT TAGGTATTGT ACCGAGAAAC

1351 CCACAACATAT CTCTATTGGC TATATGCCAA TACTCTGTCC TTCAGAGACT  
GGTGTTGATA GAGATAACCG ATATACGGTT ATGAGACAGG AAGTCTCTGA

1401 GACACGGACT CTGTATTTTT ACAGGATGGG GTCCATTTAT TATTTACAAA  
CTGTGCCTGA GACATAAAAA TGTCTTACCC CAGGTAAATA ATAAATGTTT

1451 TTCACATATA CAACAACGCC GTCCCCCGTG CCCGCAGTTT TTATTAAACA  
AAGTGTATAT GTTGTGCGG CAGGGGGCAC GGGCGTCAA AATAATTGT

1501 TAGCGTGGA TCTCCGACAT CTCGGGTACG TGTTCCGGAC ATGGGCTCTT  
ATCGCACCT AGAGGCTGTA GAGCCCATGC ACAAGGCCTG TACCCGAGAA

1551 CTCCGGTAGC GCGGGAGCTT CCACATCCGA GCCCTGGTCC CATCCGTCCA  
GAGGCCATCG CCGCCTCGAA GGTGTAGGCT CGGGACCAGG GTAGGCAGGT

1601 GCGGCTCATG GTCGCTCGGC AGCTCCTTGC TCCTAACAGT GGAGGCCAGA  
CGCCGAGTAC CAGCGAGCCG TCGAGGAACG AGGATTGTCA CCTCCGGTCT

1651 CTTAGGCACA GCACAATGCC CACCACCACC AGTGTGCCGC ACAAGGCCGT  
GAATCCGTGT CGTGTACGG GTGGTGGTGG TCACACGGCG GTTCCGGCA

1701 GGCGGTAGGG TATGTGTCTG AAAATGAGCT CGGAGATTGG GCTCGCACCT  
CCGCCATCCC ATACACAGAC TTTACTCGA GCCTCTAACC CGAGCGTGGA

1751 GGACGCAGAT GGAAGACTTA AGGCAGCGGC AGAAGAAGAT GCAGGCAGCT  
CCTGCGTCTA CCTTCTGAAT TCCGTCGCCG TCTTCTTCTA CGTCCGTCTGA

1801 GAGTTGTTGT ATTCTGATAA GAGTCAGAGG TAACTCCCGT TGCGGTGCTG  
CTCAACAACA TAAGACTATT CTCAGTCTCC ATTGAGGGCA ACGCCACGAC

FIG. 9B



pCMV-NS34A

1851 TTAACGGTGG AGGGCAGTGT AGTCTGAGCA GTACTCGTTG CTGCCGCGCG  
AATTGCCACC TCCCGTCACA TCAGACTCGT CATGAGCAAC GACGGCGCGC

---

1901 CGCCACCAGA CATAATAGCT GACAGACTAA CAGACTGTTC CTTTCCATGG  
GCGGTGGTCT GTATTATCGA CTGTCTGATT GTCTGACAAG GAAAGGTACC

---

+2 M A P  
EcoRI  
-----

1951 GTCTTTTCTG CAGTCACCGT CGTCGACCTA AGAATTCAAC ATGGCGCCCA  
CAGAAAAGAC GTCAGTGGCA GCAGCTGGAT TCTTAAGTGG TACCGCGGGT

---

+2 I T A Y A Q Q T R G L L G C I I T  
2001 TCACGGCGTA CGCCCAGCAG ACAAGGGGCC TCCTAGGGTG CATAATCACC  
AGTGCCGCAT GCGGGTCGTC TGTTCGCCG AGGATCCAC GTATTAGTGG

---

+2 S L T G R D K N Q V E G E V Q I V  
2051 AGCCTAACTG GCCGGGACAA AAACCAAGTG GAGGGTGAGG TCCAGATTGT  
TCGGATTGAC CGGCCCTGTT TTTGGTTTAC CTCCCACTCC AGGTCTAACA

---

+2 S T A A Q T F L A T C I N G V C  
2101 GTCAACTGCT GCCCAAACCT TCCTGGCAAC GTGCATCAAT GGGGTGTGCT  
CAGTTGAEGA CGGGTTTGA AGGACCGTTG CACGTAGTTA CCCACACGA

---

+2 W T V Y H G A G T R T I A S P K G  
2151 GGACTGTCTA CCACGGGGCC GGAACGAGGA CCATCGCGTC ACCCAAGGGT  
CCTGACAGAT GGTGCCCCGG CTTGCTCCT GGTAGCGCAG TGGGTTCCTCA

---

-2 P V I Q M Y T N V D Q D L V G W P  
2201 CCTGTCTATC AGATGTATAC CAATGTAGAC CAAGACCTTG TGGGCTGGCC  
GGACAGTAGG TCTACATATG GTTACATCTG GTTCTGGAAC ACCCGACCGG

---

+2 A S Q G T R S L T P C T C G S S  
2251 CGCTTCGCAA GGTACCCGCT CATTGACACC CTGCACTTGC GGCTCCTCGG  
GCGAAGCGTT CCATGGGCGA GTAACGTGTT GACGTGAAC CCGAGGAGCC

---

+2 D L Y L V T R H A D V I P V R R R  
2301 ACCTTTACCT GGTACAGAGG CACGCCGATG TCATTCCCGT GCGCCGGCGG  
TGGAATGGA CCAGTGCTCC GTGCGGCTAC AGTAAGGGCA CGCGGCCGCC

---

+2 G D S R G S L L S P R P I S Y L K  
2351 GGTGATAGCA GGGGCAGCCT GCTGTCGCCC CGGCCCATTT CCTACTTGAA  
CCTATATCGT CCCCCTCGGA CGACAGCGGG GCGGGGTAAA GGATGAACCT

---

+2 G S S G G P L L C P A G H A V G  
2401 AGGCTCCTCG GGGGGTCCGC TGTGTGCCCC CGCGGGGCAC GCCGTGGGCA  
TCCGAGGAGC CCCCAGGCG ACAACACGGG GCGCCCCGTG CGGCACCCGT

---

+2 I F R A A V C T R G V A K A V D F  
2451 TATTTAGGGC CGCGGTGTGC ACCCGTGGAG TGGCTAAGGC GGTGGACTTT  
ATAAATCCCG GCGCCACACG TGGGCACCTC ACCGATTCCG CCACCTGAAA

---

+2 I P V E N L E T T M R S P V F T D  
2501 ATCCCTGTGG AGAACCTAGA GACAACCATG AGGTCCCCGG TGTTCACGGA  
TAGGGACACC TCTTGGATCT CTGTTGGTAC TCCAGGGGCC ACAAGTGCCT

FIG. 9C



pCMV-NS34A

+2 N S S P P V V P Q S F Q V A H L  
2551 TAACTCCTCT CCACCAGTAG TGCCCCAGAG CTTCCAGGTG GCTCACCTCC  
ATTGAGGAGA GGTGGTCATC ACGGGGTCTC GAAGGTCCAC CGAGTGGAGG

---

+2 H A P T G S G K S T K V P A A Y A  
2601 ATGCTCCAC AGGCAGCGC AAAAGCACCA AGGTCCCGC TGCATATGCA  
TACGAGGGTG TCCGTCGCCG TTTTCGTGGT TCCAGGGCCG ACGTATACGT

---

+2 A Q G Y K V L V L N P S V A A T L  
2651 GCTCAGGGCT ATAAGGTGCT AGTACTCAAC CCCTCTGTTG CTGCAACACT  
CGAGTCCCGA TATTCCACGA TCATGAGTTG GGGAGACAAC GACGTTGTGA

---

+2 G F G A Y M S K A H G I D P N I  
2701 GGGCTTTGGT GCTTACATGT CCAAGGCTCA TGGGATCGAT CCTAACATCA  
CCCGAAACCA CGAATGTACA GGTTCGAGT ACCCTAGCTA GGATTGTAGT

---

+2 R T G V R T I T T G S P I T Y S T  
2751 GGACCGGGGT GAGAACAATT ACCACTGGCA GCCCATCAC GTACTCCACC  
CCTGGCCCCA CTCTTGTTAA TGGTGACCGT CGGGGTAGTG CATGAGGTGG

---

+2 Y G K F L A D G G C S G G A Y D I  
2801 TACGGCAAGT TCCTTGCCGA CGGCGGGTGC TCGGGGGGCG CTTATGACAT  
ATGCCGTTC AAGAACGGCT GCCGCCACG AGCCCCCGC GAATACTGTA

---

+2 I I C D E C H S T D A T S I L G  
2851 AATAATTGT GACGAGTGCC ACTCCACGGA TGCCACATCC ATCTTGGGCA  
TTATTAAACA CTGCTCACGG TGAGGTGCCT ACGGTGTAGG TAGAACCGT

---

+2 I G T V L D Q A E T A G A R L V V  
2901 TTGGCACTGT CCTTGACCAA GCAGAGACTG CGGGGGCGAG ACTGGTTGTG  
AACCGTGACA GGAAGTGGT CGTCTCTGAC GCGGGCGCTC TGACCAACAC

---

+2 L A T A T P P G S V T V P H P N I  
2951 CTCGCCACCG CCACCCCTCC GGGCTCCGTC ACTGTGCCCC ATCCCAACAT  
GAGCGGTGGC GGTGGGGAGG CCCGAGGCAG TGACACGGGG TAGGGTTGTA

---

+2 E E V A L S T T G E I P F Y G K  
3001 CGAGGAGGTT GCTCTGTCCA CCACCGGAGA GATCCCTTTT TACGGCAAGG  
GCTCCTCCAA CGAGACAGGT GGTGGCCTCT CTAGGGAAAA ATGCCGTTCC

---

+2 A I P L E V I K G G R H L I F C H  
3051 CTATCCCCCT CGAAGTAATC AAGGGGGGGA GACATCTCAT CTTCTGTGAT  
GATAGGGGGA GCTTCATTAG TTCCCCCCT CTGTAGAGTA GAAGACAGTA

---

+2 S K K K C D E L A A K L V A L G I  
3101 TCAAAGAAGA AGTGCGACGA ACTCGCGCA AAGCTGGTCG CATTGGGCAT  
AGTTTCTTCT TCACGCTGCT TGAGCGGCGT TTCGACCAGC GTAACCCGTA

---

+2 N A V A Y Y R G L D V S V I P T  
3151 CAATGCCGTG GCCTACTACC GCGGTCTTGA CGTGTCCGTC ATCCCGACCA  
GTTACGGCAC CGGATGATGG CGCCAGAACT GCACAGGCAG TAGGGCTGGT

---

+2 S G D V V V V A T D A L M T G Y T  
3201 GCGGCGATGT TGTGCTGCTG GCAACCGATG CCCTCATGAC CGGCTATACC  
CGCCGCTACA ACAGCAGCAC CGTTGGCTAC GGGAGTACTG GCCGATATGG

FIG. 9D



pCMV-NS34A

+2 G D F D S V I D C N T C V T Q T V  
3251 GCGGACTTCG ACTCGGTGAT AGACTGCAAT ACGTGTGTCA CCCAGACAGT  
CCGCTGAAGC TGAGCCACTA TCTGACGTTA TGCACACAGT GGGTCTGTCA

---

+2 D F S L D P T F T I E T I T L P  
3301 CGATTTTCAGC CTTGACCCTA CCTTCACCAT TGAGACAATC ACGCTCCCC  
GCTAAAGTCG GAACTGGGAT GGAAGTGGTA ACTCTGTTAG TGCGAGGGGG

---

+2 Q D A V S R T Q R R G R T G R G K  
3351 AAGATGCTGT CTCCCGCACT CAACGTCGGG GCAGGACTGG CAGGGGGAAG  
TTCTACGAGA GAGGGCGTGA GTTGACAGCC CGTCCTGACC GTCCCCCTTC

---

+2 P G I Y R F V A P G E R P S G M F  
3401 CCAGGCATCT ACAGATTTGT GGCACCGGGG GAGCGCCCT CCGGCATGTT  
GGTCCGTAGA TGTCTAAACA CCGTGGCCCC CTCGCGGGGA GGCCGTACAA

---

+2 D S S V L C E C Y D A G C A W Y  
3451 CGACTCGTCC GTCCTCTGTG AGTGCTATGA CGCAGGCTGT GCTTGGTATG  
GCTGAGCAGG CAGGAGACAC TCACGATACT GCGTCCGACA CGAACCATAC

---

+2 E L T P A E T T V R L R A Y M N T  
3501 AGCTCACGCC CGCCGAGACT ACAGTTAGGC TACGAGCGTA CATGAACACC  
TCGAGTGCGG GCGGCTCTGA TGTCAATCCG ATGCTCGCAT GTACTTGTGG

---

+2 P G L P V C Q D H L E F W E G V F  
3551 CCGGGGCTTC CCGTGTGCCA GGACCATCTT GAATTTTGGG AGGGCGTCTT  
GGCCCCGAAG GGCACACGGT CCTGGTAGAA CTTAAACCC TCCCGCAGAA

---

+2 T G L T H I D A H F L S Q T K Q  
StuI  
-----  
3601 TACAGGCCTC ACTCATATAG ATGCCCACTT TCTATCCAG ACAAGCAGA  
ATGTCCGGAG TGAGTATATC TACGGGTGAA AGATAGGGTC TGTTCGTCT

---

+2 S G E N L P Y L V A Y Q A T V C A  
3651 GTGGGGAGAA CTTTCCTTAC CTGGTAGCGT ACCAAGCCAC CGTGTGCGCT  
CACCCTCTT GGAAGGAATG GACCATCGCA TGGTTCGGTG GCACACGCGA

---

+2 R A Q A P P P S W D Q M W K C L I  
3701 AGGGCTCAAG CCCCTCCCC ATCGTGGGAC CAGATGTGGA AGTGTTTGAT  
TCCGAGTTC GGGGAGGGGG TAGCACCTG GTCTACACCT TCACAACTA

---

+2 R L K P T L H G P T P L L Y R L  
3751 TCGCCTCAAG CCCACCCTCC ATGGGCCAAC ACCCCTGCTA TACAGACTGG  
AGCGGAGTTC GGGTGGGAGG TACCGGTTG TGGGGACGAT ATGTCTGACC

---

+2 G A V Q N E I T L T H P V T K Y I  
3801 GCGCTGTTCA GAATGAAATC ACCCTGACGC ACCCAGTCAC CAAATACATC  
CGCGACAAGT CTTACTTTAG TGGGACTGCG TGGGTCAGTG GTTTATGTAG

---

+2 M T C M S A D L E V V T S T W V L  
3851 ATGACATGCA TGTCGGCCGA CCTGGAGGTC GTCACGAGCA CCTGGGTGCT  
TACTGTACGT ACAGCCGGCT GGACCTCCAG CAGTGCTCGT GGACCCACGA

---

+2 V G G V L A A L A A Y C L S T G  
3901 CGTTGGCGGC GTCCTGGCTG CTTTGGCCGC GTATTGCCTG TCAACAGGCT  
GCAACCGCCG CAGGACCGAC GAAACCGGCG CATAACGGAC AGTTGTCCGA

FIG. 9E



pCMV-NS34A

```

+2 C V V I V G R V V L S G K P A I I
3951 GCGTGGTCAT AGTGGGCAGG GTCGTCTTGT CCGGGAAGCC GGCAATCATA
    CGCACCAGTA TCACCCGTCC CAGCAGAACA GGCCCTTCGG CCGTTAGTAT

+2 P D R E V L Y R E F D E M E E C
4001 CCTGACAGGG AAGTCCTCTA CCGAGAGTTC GATGAGATGG AAGAGTGCTA
    GGACTGTCCC TTCAGGAGAT GGCTCTCAAG CTACTCTACC TTCTCACGAT

      BamHI      MluI
      -----
4051 GGATCCACTA CGCGTTAGAG CTCGCTGATC AGCCTCGACT GTGCCTTCTA
    CCTAGGTGAT GCGCAATCTC GAGCGACTAG TCGGAGCTGA CACGGAAGAT

4101 GTTGCCAGCC ATCTGTTGTT TGCCCCCTCCC CCGTGCCTTC CTTGACCCTG
    CAACGGTCGG TAGACAACAA ACGGGGAGGG GGCACGGAAG GAACTGGGAC

4151 GAAGGTGCCA CTCCCCTGT CTTTCTCTAA TAAAATGAGG AAATTGCATC
    CTTCCACGGT GAGGGTGACA GGAAAGGATT ATTTTACTCC TTTAACGTAG

4201 GCATTGCTG AGTAGGTGTC ATTCTATTCT GGGGGGTGGG GTGGGGCAGG
    CGTAACAGAC TCATCCACAG TAAGATAAGA CCCCCACCC CACCCCGTCC

4251 ACAGCAAGGG GGAGGATTGG GAAGACAATA GCAGGCATGC TGGGGAGCTC
    TGTCGTTCCC CCTCCTAACC CTTCTGTAT CGTCCGTACG ACCCCTCGAG

4301 TTCCGCTTCC TCGCTCACTG ACTCGCTGCG CTCGGTCTGT CGGCTGCGGC
    AAGGCGAAGG AGCGAGTGAC TGAGCGACGC GAGCCAGCAA GCCGACGCCG

4351 GAGCGGTATC AGCTCACTCA AAGGCGGTAA TACGGTTATC CACAGAATCA
    CTCGCCATAG TCGAGTGAGT TTCCGCCATT ATGCCAATAG GTGTCTTAGT

4401 GGGGATAACG CAGGAAAGAA CATGTGAGCA AAAGGCCAGC AAAAGGCCAG
    CCCCTATTGC GTCCCTTCTT GTACACTCGT TTTCCGGTCG TTTTCCGGTC

4451 GAACCGTAAA AAGGCCGCGT TGCTGGCGTT TTTCCATAGG CTCCGCCCCC
    CTTGGCATT TTCCGGCGCA ACGACCGCAA AAAGGTATCC GAGGCGGGGG

4501 CTGACGAGCA TCACAAAAAT CGACGCTCAA GTCAGAGGTG GCGAAACCCG
    GACTGCTCGT AGTGTTTTTA GCTGCGAGTT CAGTCTCCAC CGCTTTGGGC

4551 ACAGGACTAT AAAGATACCA GCGTTTTCCC CCTGGAAGCT CCTCGTGCG
    TGTCCTGATA TTTCTATGGT CCGCAAAGGG GGACCTTCGA GGGAGCACGC

4601 CTCTCCTGTT CCGACCCTGC CGCTTACCGG ATACCTGTCC GCCTTTCTCC
    GAGAGGACAA GGCTGGGACG GCGAATGGCC TATGGACAGG CGGAAAGAGG

4651 CTTGCGGAAG CGTGGCGCTT TCTCAATGCT CACGCTGTAG GTATCTCAGT
    GAAGCCCTTC GCACCGCGAA AGAGTTACGA GTGCGACATC CATAGAGTCA

4701 TCGGTGTAGG TCGTTTCGCT CAAGCTGGGC TGTGTGCACG AACCCCCCGT
    AGCCACATCC AGCAAGCGAG GTTCGACCCG ACACACGTGC TTGGGGGGCA

4751 TCAGCCCGAC CGCTGCGCCT TATCCGGTAA CTATCGTCTT GAGTCCAACC
    AGTCGGGCTG GCGACGCGGA ATAGGCCATT GATAGCAGAA CTCAGGTTGG

4801 CGGTAAGACA CGACTTATCG CCACTGGCAG CAGCCACTGG TAACAGGATT
    GCCATTCTGT GCTGAATAGC GGTGACCGTC GTCGGTGACC ATTGTCTTAA
  
```

FIG. 9F



pCMV-NS34A

4851 AGCAGAGCGA GGTATGTAGG CGGTGCTACA GAGTTCTTGA AGTGGTGGCC  
TCGTCTCGCT CCATACATCC GCCACGATGT CTCAAGAACT TCACCACCGG

4901 TAACTACGGC TACACTAGAA GGACAGTATT TGGTATCTGC GCTCTGCTGA  
ATTGATGCCG ATGTGATCTT CCTGTCATAA ACCATAGACG CGAGACGACT

4951 AGCCAGTTAC CTTTCGGA AAAA AGAGTTGGTA GCTCTTGATC CGGCAAACAA  
TCGGTCAATG GAAGCCTTTT TCTCAACCAT CGAGAAGTAG GCCGTTTGT

5001 ACCACCGCTG GTAGCGGTGG TTTTTTTGTT TGCAAGCAGC AGATTACGCG  
TGGTGGCGAC CATCGCCACC AAAAAACAA ACGTTCGTCG TCTAATGCGC

5051 CAGAAAAAAA GGATCTCAAG AAGATCCTTT GATCTTTTCT ACGGGGTCTG  
GTCTTTTTTT CCTAGAGTTC TTCTAGGAAA CTAGAAAAGA TGCCCCAGAC

5101 ACGCTCAGTG GAACGAAAAC TCACGTTAAG GGATTTTGGT CATGAGATTA  
TGCGAGTCAC CTTGCTTTTG AGTGCAATTC CCTAAAACCA GTACTCTAAT

5151 TCAAAAAGGA TCTTCACCTA GATCCTTTTA AATTAAAAAT GAAGTTTTAA  
AGTTTTTCCT AGAAGTGGAT CTAGGAAAAT TTAATTTTAA CTTCAAAATT

5201 ATCAATCTAA AGTATATATG AGTAACTTG GTCTGACAGT TACCAATGCT  
TAGTTAGATT TCATATATAC TCATTTGAAC CAGACTGTCA ATGGTTACGA

5251 TAATCAGTGA GGCACCTATC TCAGCGATCT GTCTATTTTCG TTCATCCATA  
ATTAGTCACT CCGTGGATAG AGTCGCTAGA CAGATAAAGC AAGTAGGTAT

5301 GTTGCCCTGAC TCCCCGTCGT GTAGATAACT ACGATACGGG AGGGCTTACC  
CAACGGACTG AGGGGCAGCA CATCTATTGA TGCTATGCCC TCCCGAATGG

5351 ATCTGGCCCC AGTGCTGCAA TGATACCGCG AGACCCACGC TCACCGGCTC  
TAGACCGGGG TCACGACGTT ACTATGGCGC TCTGGGTGCG AGTGGCCGAG

5401 CAGATTTATC AGCAATAAAC CAGCCAGCCG GAAGGGCCGA GCGCAGAAGT  
GTCTAAATAG TCGTTATTG GTCGGTGCGC CTTCCCGGCT CGCGTCTTCA

5451 GGTCCTGCAA CTTTATCCGC CTCCATCCAG TCTATTAATT GTTGCCGGGA  
CCAGGACGTT GAAATAGGCG GAGGTAGGTC AGATAATTAA CAACGGCCCT

5501 AGCTAGAGTA AGTAGTTCGC CAGTTAATAG TTTGCGCAAC GTTGTTGCCA  
TCGATCTCAT TCATCAAGCG GTCAATTATC AAACGCGTTG CAACAACGGT

5551 TTGCTACAGG CATCGTGGTG TCACGCTCGT CGTTTGGTAT GGCTTCATTC  
AACGATGTCC GTAGCACCAC AGTGCGAGCA GCAAACCATA CCGAAGTAAG

5601 AGCTCCGGTT CCCAACGATC AAGGCGAGTT ACATGATCCC CCATGTTGTG  
TCGAGGCCAA GGGTTGCTAG TTCCGCTCAA TGTACTAGGG GGTACAACAC

5651 CAAAAAGCG GTTAGCTCCT TCGGTCTCTC GATCGTTGTC AGAAGTAAGT  
GTTTTTCGC CAATCGAGGA AGCCAGGAGG CTAGCAACAG TCTTCATTCA

5701 TGGCCGCAGT GTTATCACTC ATGGTTATGG CAGCACTGCA TAATTCTCTT  
ACCGGCGTCA CAATAGTGAG TACCAATACC GTCGTGACGT ATTAAGAGAA

5751 ACTGTCATGC CATCCGTAAG ATGCTTTTCT GTGACTGGTG AGTACTCAAC  
TGACAGTACG GTAGGCATTG TACGAAAAGA CACTGACCAC TCATGAGTTG

FIG. 9G



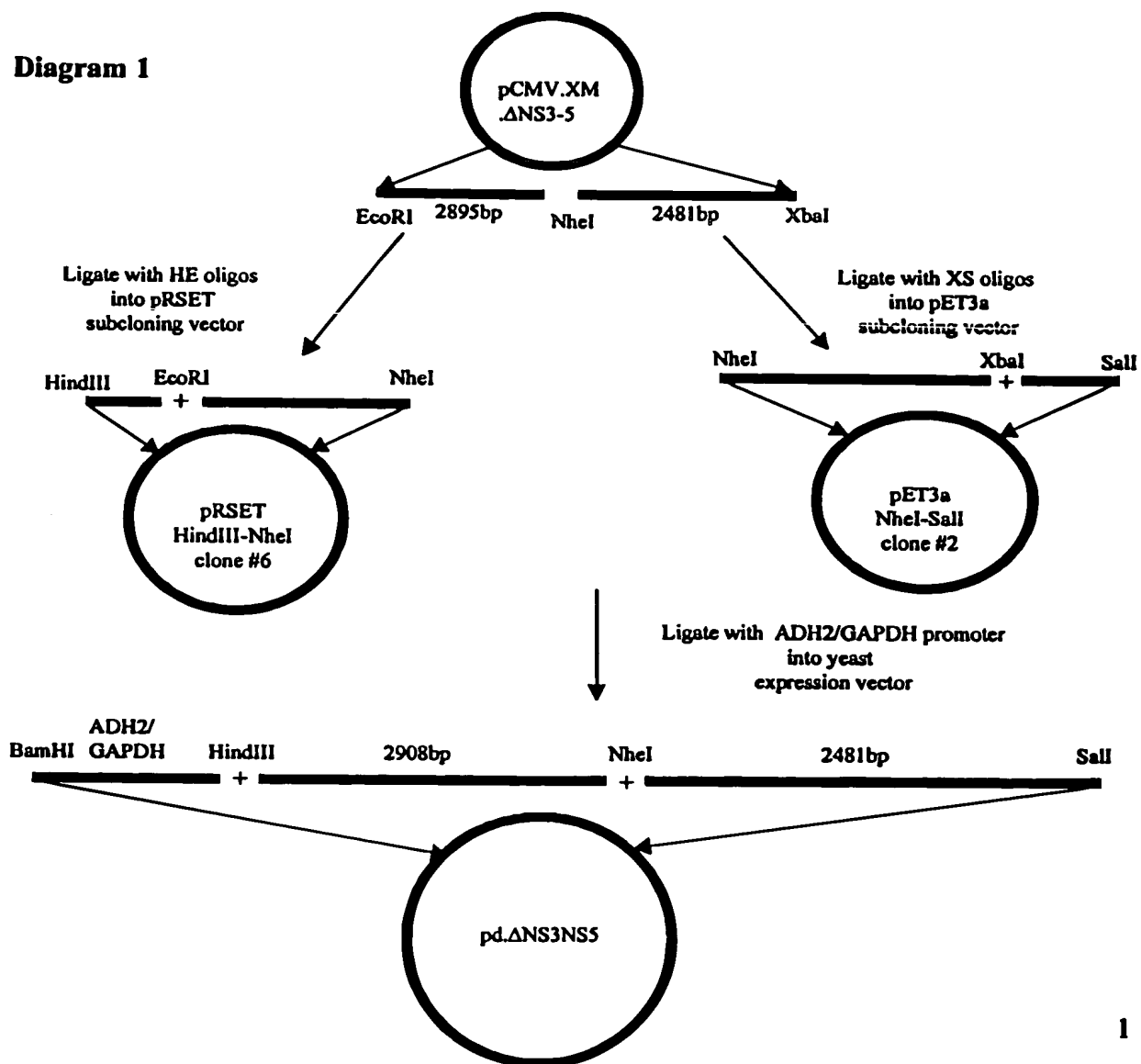
pCMV-NS34A

5801	CAAGTCATTC	TGAGAATAGT	GTATGCGGCG	ACCGAGTTGC	TCTTGCCCGG
	GTTCAAGTAAG	ACTCTTATCA	CATACGCCGC	TGGCTCAACG	AGAACGGGCC
5851	CGTCAATACG	GGATAATACC	GCGCCACATA	GCAGAACTTT	AAAAGTGCTC
	GCAGTTATGC	CCTATTATGG	CGCGGTGTAT	CGTCTTGAAA	TTTTCACGAG
5901	ATCATTGGAA	AACGTTCTTC	GGGGCGAAAA	CTCTCAAGGA	TCTTACCGCT
	TAGTAACCTT	TTGCAAGAAG	CCCCGCTTTT	GAGAGTTCCT	AGAATGGCGA
5951	GTTGAGATCC	AGTTCGATGT	AACCCACTCG	TGCACCCAAC	TGATCTTCAG
	CAACTCTAGG	TCAAGCTACA	TTGGGTGAGC	ACGTGGGTTG	ACTAGAAGTC
6001	CATCTTTTAC	TTTCACCAGC	GTTTCTGGGT	GAGCAAAAAC	AGGAAGGCAA
	GTAGAAAATG	AAAGTGGTCG	CAAAGACCCA	CTCGTTTTTG	TCCTTCCGTT
6051	AATGCCGCAA	AAAAGGGAAT	AAGGGCGACA	CGGAAATGTT	GAATACTCAT
	TTACGGCGTT	TTTTCCCTTA	TTCCCGCTGT	GCCTTTACAA	CTTATGAGTA
6101	ACTCTTCCTT	TTTCAATATT	ATTGAAGCAT	TTATCAGGGT	TATTGTCTCA
	TGAGAAGGAA	AAAGTTATAA	TAACTTCGTA	AATAGTCCCA	ATAACAGAGT
6151	TGAGCGGATA	CATATTTGAA	TGTATTTAGA	AAAATAAACA	AATAGGGGTT
	ACTCGCCTAT	GTATAAACTT	ACATAAATCT	TTTTATTTGT	TTATCCCCAA
6201	CCGCGCACAT	TTCCCCGAAA	AGTGCCACCT	GACGTCTAAG	AAACCATTAT
	GGCGCGTGTA	AAGGGGCTTT	TCACGGTGGA	CTGCAGATTC	TTTGGAATA
6251	TATCATGACA	TTAACCTATA	AAAATAGGCG	TATCACGAGG	CCCTTTCGTC
	ATAGTACTGT	AATTGGATAT	TTTTATCCGC	ATAGTGCTCC	GGGAAAGCAG

FIG. 9H



**Diagram 1**

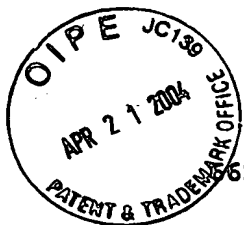


**FIG. 10**



MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuVal  
2 AGCTTACAAAACAAATTACCATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTA  
TCGAATGTTTTGTTAAGTGGTACCGACGTATACGTGAGTCCCGATATTCCACGATCAT  
^ ^ ^ ^  
1 HIND3, 21 NCOI, 30 NDEI, 58 SCAI,  
  
LeuAsnProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGly  
62 CTCAACCCCTCTGTTGCTGCAACACTGGGCTTTGGTGCTTACATGTCCAAGGCTCATGGG  
GAGTTGGGGAGACAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCC  
  
IleAspProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyr  
122 ATCGATCCTAACATCAGGACCGGGGTGAGAACAATTACCACTGGCAGCCCCATCACGTAC  
TAGCTAGGATTGTAGTCCTGGCCCCACTCTTGTTAATGGTGACCGTCGGGGTAGTGCATG  
^  
122 CLAI,  
  
SerThrTyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIle  
182 TCCACCTACGGCAAGTTCCCTTGCCGACGGCGGGTGCTCGGGGGGCGCTTATGACATAATA  
AGGTGGATGCCGTTCAAGGAACGGCTGCCGCCACGAGCCCCCGGAATACTGTATTAT  
  
IleCysAspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeu  
242 ATTTGTGACGAGTGCCACTCCACGGATGCCACATCCATCTTGGGCATTGGCACTGTCCTT  
TAAACACTGCTCACGGTGAGGTGCCTACGGTGTAGGTAGAACCCGTAACCGTGACAGGAA  
  
AspGlnAlaGluThrAlaGlyAlaArgLeuValValLeuAlaThrAlaThrProProGly  
302 GACCAAGCAGAGACTGCGGGGGCGAGACTGGTTGTGCTCGCCACCGCCACCCCTCCGGGC  
CTGGTTCGTCTCTGACGCCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCCG  
^  
309 ALWN1,  
  
SerValThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIle  
362 TCCGTCACTGTGCCCCATCCCAACATCGAGGAGGTTGCTCTGTCCACCACCGGAGAGATC  
AGGCAGTGACACGGGTAGGGTTGTAGCTCCTCCAACGAGACAGGTGGTGGCCTCTCTAG  
  
ProPheTyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePhe  
422 CCTTTTACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGAGACATCTCATCTTC  
GGAAAAATGCCGTTCCGATAGGGGGAGCTTCATTAGTTCCCCCCTCTGTAGAGTAGAAG  
  
CysHisSerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsn  
482 TGTCATTCAAAGAAGAAGTGCGACGAACCTCGCCGCAAAGCTGGTCGCATTGGGCATCAAT  
ACAGTAAGTTTCTTCTTCACGCTGCTTGAGCGGCGTTTCGACCAGCGTAACCCGTAGTTA  
  
AlaValAlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValVal  
542 GCCGTGGCCTACTACCGCGGTCTTGACGTGTCCGTATCCCGACCAGCGCGATGTTGTC  
CGGCACCGGATGATGGCGCCAGAATGCACAGGCAGTAGGGCTGGTCGCCGCTACAACAG  
^ ^  
556 SAC2, 566 DRD1,  
  
ValValAlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAsp  
602 GTCGTGGCAACCGATGCCCTCATGACCGGCTATACCGGCGACTTCGACTCGGTGATAGAC  
CAGCACC GTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTG  
^  
621 BSPH1,  
  
CysAsnThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGlu

FIG. 11A



62 TGCAATACGTGTGTCAACCAGACAGTCGATTTAGCCTTGACCCTACCTTCACCATTGAG  
ACGTTATGCACACAGTGGGTCTGTGAGCTAAAGTCGGAACGGGATGGAAGTGGTAAC

722 ThrIleThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArg  
ACAATCACGCTCCCCAAGATGCTGTCTCCGCACTCAACGTCGGGGCAGGACTGGCAGG  
TGTTAGTGCAGGGGGTTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCTGACCGTCC

782 GlyLysProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAsp  
GGGAAGCCAGGCATCTACAGATTTGTGGCACCAGGGGAGCGCCCCCTCCGGCATGTTGAC  
CCCTTCGGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAGCTG  
^ ^

822 BGLI, 839 DRD1,

842 SerSerValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAla  
TCGTCCGTCTCTGTGAGTGCTATGACGCAGGCTGTGCTTGGTATGAGCTCACGCCCCGCC  
AGCAGGCAGGAGACACTCACGATACTGCGTCCGACACGAACCATACTCGAGTGCGGGCGG  
^

887 SACI,

902 GluThrThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAsp  
GAGACTACAGTTAGGCTACGAGCGTACATGAACACCCCGGGGCTTCCCGTGTGCCAGGAC  
CTCTGATGTCAATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCACACGGTCCGTG  
^

937 SMAI XMAI,

962 HisLeuGluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeu  
CATCTTGAATTTTGGGAGGGCGTCTTTACAGGCCTCACTCATATAGATGCCCACTTTCTA  
GTAGAACTTAAAACCCCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGAT  
^

991 STUI,

1022 SerGlnThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrVal  
TCCCAGACAAAGCAGAGTGGGGAGAACCTTCCCTTACCTGGTAGCGTACCAAGCCACCGTG  
AGGGTCTGTTTCGTCTCACCCCTCTTGGAAGGAATGGACCATCGCATGGTTCGGTGGCAC  
^

1075 DRA3,

1082 CysAlaArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysLeuIleArg  
TGCGCTAGGGCTCAAGCCCCTCCCCATCGTGGGACCAGATGTGGAAGTGTGTTGATTGCG  
ACGCGATCCCGAGTTCGGGGAGGGGGTAGCACCTGGTCTACACCTTCACAACTAAGCG

1142 LeuLysProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsn  
CTCAAGCCCACCCTCCATGGGCCAACACCCCTGCTATACAGACTGGGCGCTGTTTCAAGAT  
GAGTTCGGGTGGGAGGTACCCGGTTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTA  
^

1156 NCOI,

1202 GluIleThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeu  
GAAATCACCCCTGACGCACCCAGTCACCAAATACATCATGACATGCATGTCGGCCGACCTG  
CTTTAGTGGGACTGCGTGGGTGAGTGGTTTATGTAGTACTGTACGTACAGCCGGCTGGAC  
^ ^ ^ ^ ^

1236 BSPH1, 1240 DRD1, 1243 AVA3, 1251 EAG1 XMA3, 1256 DRD1,

1262 GluValValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyr  
GAGGTCGTCACGAGCACCTGGGTGCTCGTTGGCGGCGTCCTGGCTGCTTTGGCCGCGTAT  
CTCCAGCAGTGCTCGTGACCCACGAGCAACCGCCGAGGACCGACGAAACCGGCGCATA

FIG. 11B



1322 CysLeuSerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAla  
TGCTGTCAACAGGCTGCGTGGTCATAGTGGGCAGGGTCGTCTTGTCCGGAAGCCGGCA  
ACGGACAGTTGTCCGACGCACCAGTATCACCCGTCCCAGCAGAACAGGCCCTTCGGCCGT  
^

1375 NAEI,

1382 IleIleProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGln  
ATCATACCTGACAGGGAAGTCCCTCTACCGAGAGTTTCGATGAGATGGAAGAGTGCTCTCAG  
TAGTATGGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCACGAGAGTC  
^

1391 DRD1,

1442 HisLeuProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeu  
CACTTACCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGCAGAAGGCCCTC  
GTGAATGGCATGTAGCTCGTTCCTACTACGAGCGGCTCGTCAAGTTCGTCTTCCGGGAG

1502 GlyLeuLeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsn  
GGCCTCCTGCAGACCGCGTCCCGTCAGGCAGAGGTTATCGCCCCTGCTGTCCAGACCAAC  
CCGGAGGACGTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTG  
^ ^

1508 PSTI, 1513 TTH3I,

1562 TrpGlnLysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGln  
TGGCAAAACTCGAGACCTTCTGGGCGAAGCATATGTGGAAGTTCATCAGTGGGATACAA  
ACCGTTTTTGTAGCTCTGGAAGACCCGCTTCGTATACACCTTGAAGTAGTCACCCTATGTT  
^ ^

1571 XHOI, 1592 NDEI,

1622 TyrLeuAlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPhe  
TACTTGGCGGGCTTGTCAACGCTGCCTGGTAACCCGCCATTGCTTCATTGATGGCTTTT  
ATGAACCGCCCGAACAGTTGCGACGGACCATTGGGGCGGTAACGAAGTAACCTACCGAAAA  
^

1649 BSTE2,

1682 ThrAlaAlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGly  
ACAGCTGCTGTCAACAGCCCACTAACCCTAGCCAAACCCCTCCTCTTCAACATATTGGGG  
TGTCGACGACAGTGGTCCGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCC  
^

1683 ALWN1 PVU2,

1742 GlyTrpValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGly  
GGGTGGGTGGCTGCCCAGCTCGCCGCCCCCGGTGCCGCTACTGCCTTTGTGGGCGCTGGC  
CCCACCCACCGACGGGTGAGCGGGCGGGGGCCACGGCGATGACGGAAACACCCGCGACCG  
^

1800 ESP1,

1802 LeuAlaGlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAla  
TTAGCTGGCGCCGCCATCGGCAGTGTGGACTGGGGAAGGTCCTCATAGACATCCTTGCA  
AATCGACCGCGGGGTAGCCGTCACAACCTGACCCCTTCAGGAGTATCTGTAGGAACGT  
^

1808 KAS1 NARI,

1862 GlyTyrGlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluVal  
GGGTATGGCGCGGGCGTGGCGGGAGCTCTTGTGGCATTCAAGATCATGAGCGGTGAGGTC  
CCCATACCGCGCCCGCACC GCCCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAG  
^ ^

FIG. 11C



1884 SACI, 1905 BSPH1,  
1922 Pr SerThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuVal  
CCCTCCACGGAGGACCTGGTCAATCTACTGCCCGCCATCCTCTCGCCCGGAGCCCTCGTA  
GGGAGGTGCCTCCTGGACCAGTTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCAT  
1934 TTH3I,  
1982 ValGlyValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaVal  
GTCGGCGTGGTCTGTGCAGCAATACTGCGCCGGCACGTTGGCCCGGGCGAGGGGCGAGTG  
CAGCCGCACCAGACACGTGTTATGACGCGGCCGTGCAACCGGGCCCCGCTCCCCCGTCAC  
2010 NAEI, 2023 SMAI XMAI,  
2042 GlnTrpMetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHis  
CAGTGGATGAACCGGCTGATAGCCTTCGCCTCCCGGGGAACCATGTTTCCCCACGCAC  
GTCACCTACTTGGCCGACTATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGGTGCGTG  
2073 SMAI XMAI, 2099 DRA3,  
2102 TyrValProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrVal  
TACGTGCCGGAGAGCGATGCAGCTGCCCGCGTCACTGCCATACTCAGCAGCCTCACTGTA  
ATGCACGGCCTCTCGCTACGTGACGGGCGCAGTGACGGTATGAGTCGTGCGAGTGACAT  
2121 PVU2,  
2162 ThrGlnLeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSer  
ACCCAGCTCCTGAGGCGACTGCACCAGTGGATAAGCTCGGAGTGTACCACTCCATGCTCC  
TGGGTCGAGGACTCCGCTGACGTGGTCACCTATTCGAGCCTCACATGGTGAGGTACGAGG  
2165 ALWN1, 2170 MST2,  
2222 GlySerTrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThr  
GGTTCCTGGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCGACTTTAAGACC  
CCAAGGACCGATTCCCTGTAGACCCTGACCTATACGCTCCACAACCTCGCTGAAATTCTGG  
2226 ECON1,  
2282 TrpLeuLysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArg  
TGGCTAAAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCCTTTGTGTCTGCCAGCGC  
ACCGATTTTCGATTTCGAGTACGGTGTGACGGACCCTAGGGGAAACACAGGACGGTTCGCG  
2291 ESP1, 2306 PVU2, 2316 BAMHI,  
2342 GlyTyrLysGlyValTrpArgGlyAspGlyIleMetHisThrArgCysHisCysGlyAla  
GGGTATAAGGGGGTCTGGCGAGGGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCT  
CCCATATTCCCCCAGACCGCTCCCCTGCCGTAGTACGTGTGAGCGACGGTGACACCTCGA  
2402 GluIleThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArg  
GAGATCACTGGACATGTCAAAAACGGGACGATGAGGATCGTCGGTCCCTAGGACCTGCAGG  
CTCTAGTGACCTGTACAGTTTTTGCCCTGCTACTCCTAGCAGCCAGGATCCTGGACGTCC  
2431 BSAB1, 2447 AVR2, 2454 SSE83871, 2455 PSTI,  
2462 AsnMetTrpSerGlyThrPheProIleAsnAlaTyrThrThrGlyProCysThrProLeu  
AACATGTGGAGTGGGACCTTCCCCATTAATGCCTACACCACGGGCCCCCTGTACCCCCCTT  
TTGTACACCTCACCTGGAAGGGGTAATTACGGATGTGGTGCCCGGGGACATGGGGGGAA

FIG. 11D



- 2486 ASE1, 2503 APA1,
- 2522 ProAlaProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIle  
CCTGCGCCGAACACTACAGTTTCGCGCTATGGAGGGTGTCTGCAGAGGAATACGTGGAGATA  
GGACGCGGCTTGATGTGCAAGCGGATACCTCCACAGACGTCTCCTTATGCACCTCTAT  
^
- 2559 PSTI,
- 2582 ArgGlnValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysPro  
AGGCAGGTGGGGGACTTCCACTACGTGACGGGTATGACTACTGACAATCTTAAATGCCCC  
TCCGTCCACCCCTGAAGGTGATGCACTGCCCATACTGATGACTGTTAGAATTTACGGGC  
^
- 2600 DRA3,
- 2642 CysGlnValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPhe  
TGCCAGGTCCCATCGCCGAATTTTTTACAGAATTGGACGGGGTGCGCTACATAGGTTT  
ACGGTCCAGGGTAGCGGGCTTAAAAAGTGTCTTAACCTGCCCCACGGGATGTATCCAA
- 2702 AlaProProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGlu  
GCGCCCCCTGCAAGCCCTTGCTGCGGGAGGAGGTATCATTCAGAGTAGGACTCCACGAA  
CGCGGGGGGACGTTTCGGGAACGACGCCCTCCTCCATAGTAAGTCTCATCCTGAGGTGCTT
- 2762 TyrProValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSer  
TACCCGGTAGGGTCGCAATTACCTTGCGAGCCCCGAACCGGACGTGGCCGTGTTGACGTCC  
ATGGGCCATCCCAGCGTTAATGGAACGCTCGGGCTTGGCCTGCACCGGCACAACCTGCAGG  
^
- 2763 HGIE2, 2815 AAT2,
- 2822 MetLeuThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGly  
ATGCTCACTGATCCCTCCCATATAACAGCAGAGGCGGCCGGGCGAAGGTTGGCGAGGGGA  
TACGAGTGACTAGGGAGGGTATATTGTCGTCTCCGCCGGCCCGCTTCCAACCGCTCCCT  
^
- 2856 EAG1 XMA3,
- 2882 SerProProSerValAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuLysAla  
TCACCCCTCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCTCAAGGCA  
AGTGGGGGGAGACACCGGTCGAGGAGCCGATCGGTCGATAGGCGAGGTAGAGAGTTCCGT  
^
- 2895 BALI, 2909 NHEI,
- 2942 ThrCysThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrp  
ACTTGACCGCTAACCATGACTCCCCTGATGCTGAGCTCATAGAGGCCAACCTCCTATGG  
TGAACGTGGCGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGGTTGGAGGATACC  
^ ^
- 2972 ESP1, 2975 SACI,
- 3002 ArgGlnGluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeu  
AGGCAGGAGATGGGCGCAACATCACCAGGGTTGAGTCAGAAAACAAAGTGGTGATTCTG  
TCCGTCTCTACCCGCCGTTGTAGTGGTCCCAACTCAGTCTTTTGTTCACCACTAAGAC
- 3062 AspSerPheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGlu  
GACTCCTTCGATCCGCTTGTTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGCAGAA  
CTGAGGAAGCTAGGCGAACACCGCCTCCTCCTGCTCGCCCTCTAGAGGCATGGGCGTCTT  
^
- 3102 BGL2,

FIG. 11E



IleLeuArgLysSerArgArgPheAlaGlnAlaL uProValTrpAlaArgPr AspTyr  
ATCCTGCGGAAGTCTCGGAGATTGCGCCAGGCCCTGCCGTTTGGGCGCGGCCGACTAT  
TAGGACGCCTTCAGAGCCTCTAAGCGGGTCCGGGACGGGCAAACCCGCGCCGGCCTGATA  
3149 ALWN1, 3170 EAG1 XMA3,  
AsnProProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGly  
3182 AACCCCCGCTAGTGGAGACGTGGAAGCCCGACTACGAACACCTGTGGTCCATGGC  
TTGGGGGGCGATCACCTCTGCACCTTTTTCGGGCTGATGCTTGGTGGACACCAGGTACCG  
3223 HGIE2, 3235 NCOI,  
CysProLeuProProProLysSerProProValProProProArgLysLysArgThrVal  
3242 TGCCCGCTTCCACCTCCAAAGTCCCTCCTGTGCCTCCGCCTCGGAAGAAGCGGACGGTG  
ACGGGCGAAGGTGGAGTTTCAGGGGAGGACACGGAGGCGGAGCCTTCTTCGCCTGCCAC  
ValLeuThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGly  
3302 GTCCTCACTGAATCAACCCTATCTACTGCCTTGGCCGAGCTCGCCACCAGAAGCTTTGGC  
CAGGAGTGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGGTCTTCGAAACCG  
3338 SACI, 3352 HIND3,  
SerSerSerThrSerGlyIleThrGlyAspAsnThrThrThrSerSerGluProAlaPro  
3362 AGCTCCTCAACTTCCGGCATTACGGGCGACAATACGACAACATCCTCTGAGCCCCCCCCCT  
TCGAGGAGTTGAAGGCCGTAATGCCCGCTGTTATGCTGTTGTAGGAGACTCGGGCGGGGA  
SerGlyCysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGly  
3422 TCTGGCTGCCCCCGACTCCGACGCTGAGTCCTATTCTCCATGCCCCCCTGGAGGGG  
AGACCGACGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGGACCTCCCC  
3443 EAM11051,  
GluProGlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsn  
3482 GAGCCTGGGGATCCGGATCTTAGCGACGGGTCATGGTCAACGGTCAGTAGTGAGGCCAAC  
CTCGGACCCCTAGGCCTAGAATCGCTGCCAGTACCAGTTGCCAGTCATCACTCCGGTTG  
3490 BAMHI, 3491 BSAB1, 3493 BSPE1,  
AlaGluAspValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuValThrPro  
3542 GCGGAGGATGTCGTGTGCTGCTCAATGTCTTACTCTTGACAGGCGCACTCGTCACCCCG  
CGCCTCCTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCGCGTGAGCAGTGGGGC  
3595 DRA3,  
CysAlaAlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHis  
3602 TGCGCCGCGGAAGAACAGAACTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTCAC  
ACGCGGCGCCTTCTTGCTTTGACGGGTAGTTACGTGATTGTTGAGCAACGATGCAGTG  
3606 SAC2, 3617 ALWN1, 3661 PFLM1,  
HisAsnLeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThr  
3662 CACAATTTGGTGTATTCCACCACCTCACGCAGTGCTTGCCAAAGGCAGAAGAAAGTCACA  
GTGTTAAACCACATAAGGTGGTGGAGTCGCTCACGAACGGTTTCCGCTCTTCTTCAGTGT  
3687 DRA3,  
PheAspArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAla

FIG. 11F



3722 TTTGACAGACTGCAAGTTCTGGACAGCCATTACCAGGACGTACTCAAGGAGGTTAAAGCA  
AAACTGTCTGACGTTCAAGACCTGTGCGTAATGGTCCTGCATGAGTTCTCCAATTTCTG

3782 AlaAlaSerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrPro  
GCGGCGTCAAAAGTGAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGACGCCTGACGCCC  
CGCCGCAGTTTTCACTTCCGATTGAACGATAGGCATCTCCTTCGAACGTGCGACTGCGGG  
3822 HIND3,

3842 ProHisSerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArg  
CCACACTCAGCCAAATCCAAGTTTGGTTATGGGGCAAAAGACGTCCGTTGCCATGCCAGA  
GGTGTGAGTCGGTTTAGGTTCAAACCAATACCCCGTTTTCTGCAGGCAACGGTACGGTCT  
3881 AAT2, 3896 BGLI,

3902 LysAlaValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrPro  
AAGGCCGTAACCCACATCAACTCCGTGTGGAAAGACCTTCTGGAAGACAATGTAACACCA  
TTCCGGCATTGGGTGTAGTTGAGGCACACCTTTCTGGAAGACCTTCTGTTACATTGTGGT

3962 IleAspThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGly  
ATAGACACTACCATCATGGCTAAGAACGAGTTTTCTGCGTTACGCCTGAGAAGGGGGGT  
TATCTGTGATGGTAGTAGTACCGATTCTTGCTCCAAAAGACGCAAGTCGGAATCTTCCCCCA

4022 ArgLysProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMet  
CGTAAGCCAGCTCGTCTCATCGTGTTCCTCGATCTGGGCGTGCGCGTGTGCGAAAAGATG  
GCATTCCGTCGAGCAGAGTAGCACAAAGGGGCTAGACCCGCACGCGCACACGCTTTTCTAC

4082 AlaLeuTyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPhe  
GCTTTGTACGACGTGGTTACAAAGCTCCCTTGCCGTGATGGGAAGCTCCTACGGATTC  
CGAAACATGCTGCACCAATGTTTCGAGGGGAACCGGCACTACCCTTCGAGGATGCCTAAG

4142 GlnTyrSerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThr  
CAATACTCACCAGGACAGCGGTTGAATTCCTCGTGCAAGCGTGGAAGTCCAAGAAAACC  
GTTATGAGTGGTCTGTCGCCCAACTTAAGGAGCACGTTTCGCACCTTCAGGTTCTTTTG  
4166 ECORI,

4202 ProMetGlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIle  
CCAATGGGGTTCTCGTATGATACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATC  
GGTTACCCCAAGAGCATACTATGGGCGACGAAACTGAGGTGTCTAGTACTCTCGCTGTAG  
4235 DRD1, 4242 ALWN1,

4262 ArgThrGluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIle  
CGTACGGAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCAAGCCCGGTGGCCATC  
GCATGCCTCCTCCGTTAGATGGTTACAACACTGGAGCTGGGGGTTGGGGCGCACCGGTAG  
4307 BGLI, 4314 BALI,

4322 LysSerLeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsn  
AAGTCCCTCACCAGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAGAAC  
TTCAGGGAGTGGCTCTCCGAAATACAACCCCGGGAGAATGGTTAAGTTCCCCCTCTTG  
4351 APAI,

4382 CysGlyTyrArgArgCysArgAlaSerGlyValLeuThrThrSerCysGlyAsnThrLeu  
TGCGGCTATCGCAGGTGCCGCGGAGCGCGTACTGACAACCTAGCTGTGGTAACACCCCTC

FIG. 11G



ACGCCGATAGCGTCCACGGCGCGCTCGCCGCATGACTGTTGATCGACACCATTGTGGGAG

4442 ThrCysTyrIleLysAlaArgAlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMet  
ACTTGCTACATCAAGGCCCGGGCAGCCTGTGAGCCGCAGGGCTCCAGGACTGCACCATG  
TGAACGATGTAGTTCCGGGCGCGTCCGACAGCTCGGCGTCCCGAGGTCCTGACGTGGTAC  
^

4458 SMAI XMAI,

4502 LeuValCysGlyAspAspLeuValValIleCysGluSerAlaGlyValGlnGluAspAla  
CTCGTGTGTGGCGACGACTTAGTCGTTATCTGTGAAAGCGCGGGGGTCCAGGAGGACGCG  
GAGCACACACCGCTGCTGAATCAGCAATAGACACTTTCGCGCCCCCAGGTCCTCCTGCGC  
^ ^

4514 DRD1, 4517 TTH3I,

4562 AlaSerLeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspPro  
GCGAGCCTGAGAGCCTTCACGGAGGCTATGACCAGGTACTCCGCCCCCCTGGGGACCCC  
CGCTCGGACTCTCGGAAGTGCTCCGATACTGGTCCATGAGGCGGGGGGACCCCTGGGG

4622 ProGlnProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerValAla  
CCACAACCAGAATACGACTTGGAGCTCATAACATCATGCTCCTCCAACGTGTCAGTCGCC  
GGTGTGGTCTTATGCTGAACCTCGAGTATTGTAGTACGAGGAGGTTGCACAGTCAGCGG  
^

4643 SACI,

4682 HisAspGlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAla  
CACGACGGCGCTGGAAAGAGGGTCTACTACCTCACCCGTGACCCTACAACCCCCCTCGCG  
GTGCTGCCGCGACCTTCTCCAGATGATGGAGTGGGCACTGGGATGTTGGGGGGAGCGC  
^

4737 NRUI,

4742 ArgAlaAlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIle  
AGAGCTGCGTGGGAGACAGCAAGACACACTCCAGTCAATTCCTGGCTAGGCAACATAATC  
TCTCGACGCACCCTCTGTCGTTCTGTGTGAGGTCAGTTAAGGACCGATCCGTTGTATTAG

4802 MetPheAlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePheSerValLeu  
ATGTTTGGCCCCACACTGTGGGCGAGGATGATACTGATGACCCATTTCTTTAGCGTCCTT  
TACAAACGGGGGTGTGACACCCGCTCCTACTATGACTACTGGGTAAAGAAATCGCAGGAA  
^ ^

4812 PFLM1, 4813 DRA3,

4862 IleAlaArgAspGlnLeuGluGlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSer  
ATAGCCAGGGACAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGGCCTGCTACTCC  
TATCGGTCCCTGGTCAACTTGTCCGGGAGCTAACGCTCTAGATGCCCCGGACGATGAGG  
^

4899 BGL2,

4922 IleGluProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSer  
ATAGAACCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTTCA  
TATCTTGGTGACCTAGATGGAGGTTAGTAAGTTTCTGAGGTACCGGAGTCGCGTAAAAGT  
^

4960 NCOI,

4982 LeuHisSerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGly  
CTCCACAGTTACTCTCCAGGTGAATCAATAGGGTGGCCGCATGCCTCAGAAAACCTGGG  
GAGGTGTCAATGAGAGGTCCACTTTAGTTATCCACCGGCGTACGGAGTCTTTTGAACCC  
^

5021 SPHI, 5041 KPNI,

FIG. 11H



5042 ValProProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAla  
GTACCGCCCTTGCGAGCTTGGAGACACCGGGCCCGGAGCGTCCGCGCTAGGCTTCTGGCC  
CATGGCGGGAAACGCTCGAACCTCTGTGGCCCGGGCCTCGCAGGCGCGATCCGAAGACCGG  
5070 APAI, 5097 BALI,  
ArgGlyGlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLys  
5102 AGAGGAGGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAAG  
TCTCCTCCGTCCCGACGGTATACACCGTTCATGGAGAAGTTGACCCGTCATTCTTGTTC  
5119 NDEI,  
LeuLysLeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAla  
5162 CTCAAACCTCACTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCCGGCTGGTTACGGCT  
GAGTTTGAGTGAGGTTATCGCCGGCGACCGGTCGACCTGAACAGGCCGACCAAGTGCCGA  
5180 NOTI, 5181 EAG1 XMA3, 5188 BALI, 5192 PVU2,  
GlyTyrSerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrp  
5222 GGCTACAGCGGGGGAGACATTTATCACAGCGTGTCTCATGCCCGGCCCGCTGGATCTGG  
CCGATGTCGCCCCCTCTGTAAATAGTGTGCGACAGAGTACGGGCCGGGGCGACCTAGACC  
5246 DRA3,  
PheCysLeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgOP  
5282 TTTTGCCCTACTCCTGCTTGCTGCAGGGGTAGGCATCTACCTCCTCCCAACCGATGAAGG  
AAAACGGATGAGGACGAACGACGTCCCCATCCGTAGATGGAGGAGGGGTTGGCTACTTCC  
5301 PSTI, 5331 HGIE2,  
5342 TTGGGGTAAACACTCCGGCCTAAAAAAAAAAAAAAAAATCTAGAACCCGAGTCGAC  
AACCCCATTTGTGAGGCCGGATTTTTTTTTTTTTTTAGATCTTGGGCTCAGCTG  
5378 XBAI, 5390 SALI,

FIG. 11I

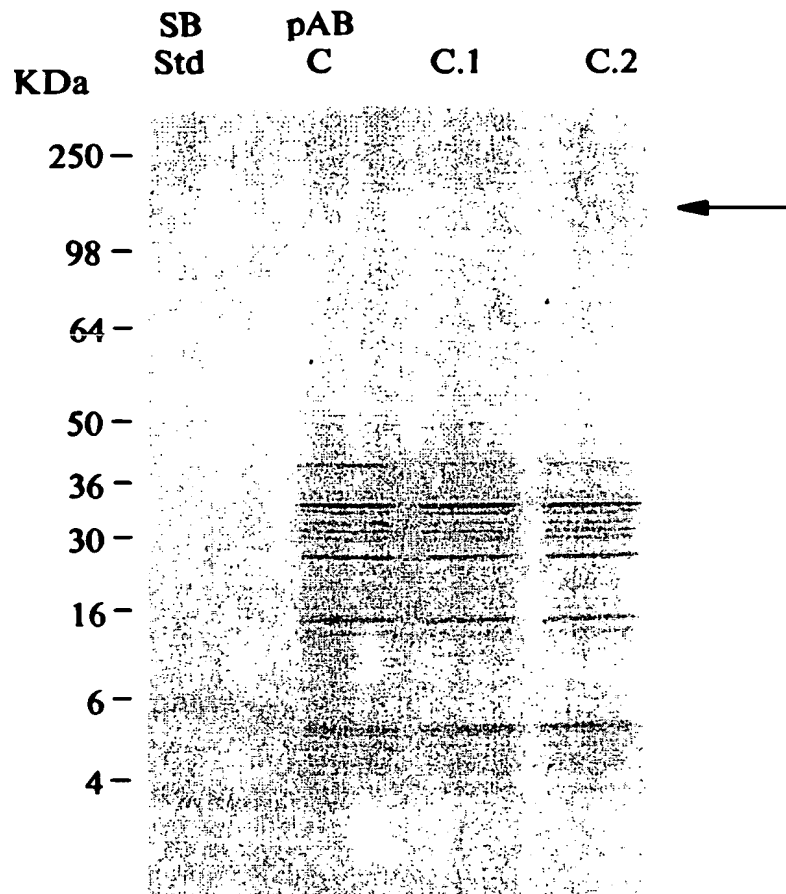


FIG. 12

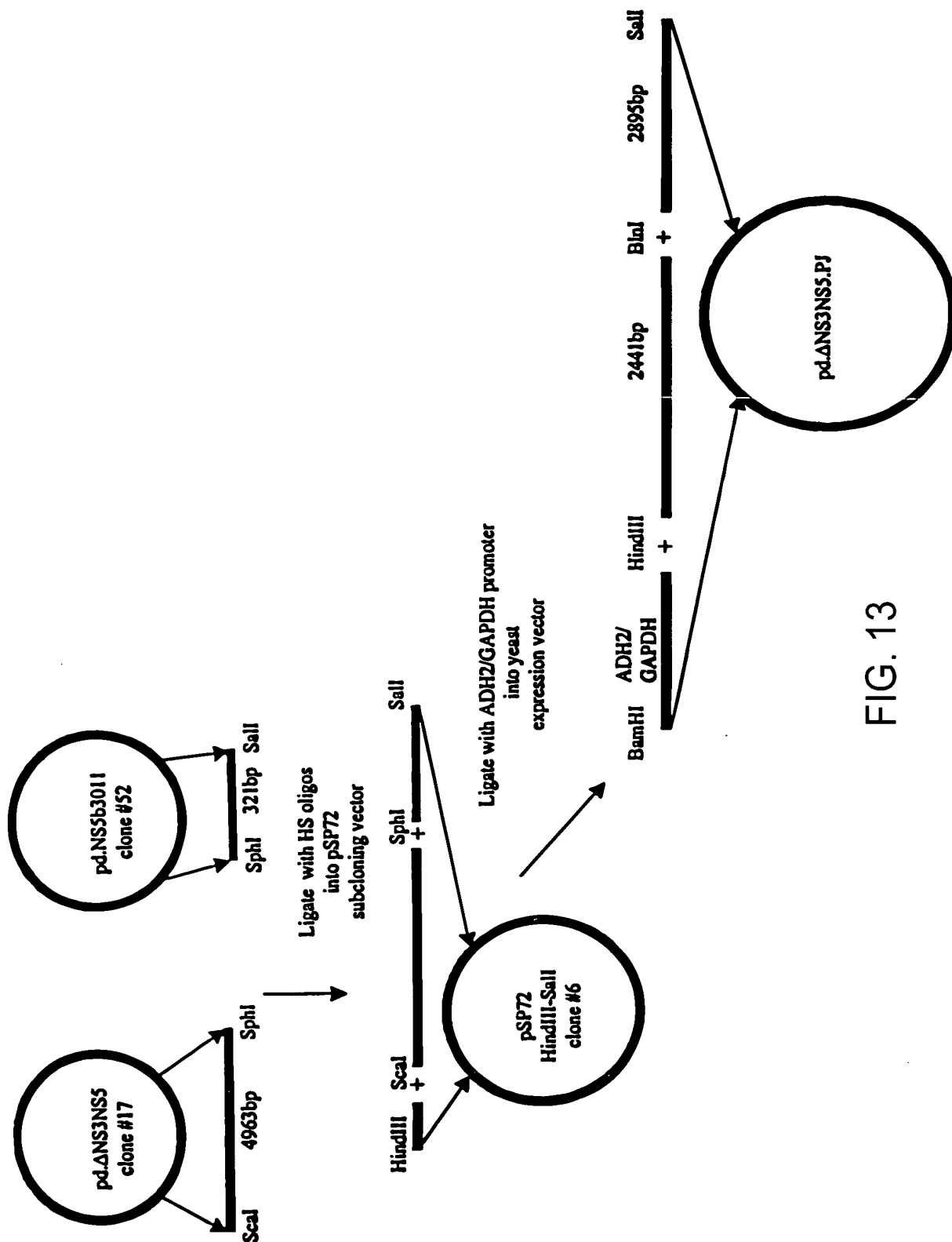


FIG. 13



MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsn  
2 AGCTTACAAAACAAATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTACTCAAC  
TCGAATGTTTTGTTTTACCGACGTATACGTCGAGTCCCGATATTCCACGATCATGAGTTG  
^ ^  
1 HIND3, 24 NDEI, 52 SCAI,  
  
ProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGlyIleAsp  
62 CCCTCTGTTGCTGCAACACTGGGCTTTGGTGCTTACATGTCCAAGGCTCATGGGATCGAT  
GGGAGACAACGACGTTGTGACCCGAAACCACGAATGTACAGGTCCGAGTACCCTAGCTA  
^  
116 CLAI,  
  
ProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyrSerThr  
122 CCTAACATCAGGACCGGGGTGAGAACAATTACCACTGGCAGCCCCATCACGTACTCCACC  
GGATTGTAGTCCTGGCCCCACTCTTGTTAATGGTGACCGTCGGGGTAGTGCATGAGGTGG  
  
TyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIleIleCys  
182 TACGGCAAGTTCCTTGCCGACGGCGGGTGCTCGGGGGGCGCTTATGACATAATAATTTGT  
ATGCCGTTCAAGGAACGGCTGCCGCCACGAGCCCCCGGAATACTGTATTATTAAACA  
  
AspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGln  
242 GACGAGTGCCACTCCACGGATGCCACATCCATCTTGGGCATTGGCACTGTCCTTGACCAA  
CTGCTCACGGTGAGGTGCCTACGGTGTAGGTAGAACCCGTAACCGTGACAGGAAC TGTT  
  
AlaGluThrAlaGlyAlaArgLeuValValLeuAlaThrAlaThrProProGlySerVal  
302 GCAGAGACTGCGGGGCGAGACTGGTTGTGCTCGCCACCGCCACCCCTCCGGGCTCCGTC  
CGTCTCTGACGCCCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCCGAGGCAG  
^  
303 ALWN1,  
  
ThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIleProPhe  
362 ACTGTGCCCCATCCCAACATCGAGGAGGTTGCTCTGTCCACCACCGGAGAGATCCCTTTT  
TGACACGGGGTAGGGTTGTAGCTCCTCCAACGAGACAGGTGGTGGCTCTCTAGGGAAAA  
  
TyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePheCysHis  
422 TACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGAGACATCTCATCTTCTGTCAT  
ATGCCGTTCCGATAGGGGGAGCTTCATTAGTTCCCCCCTCTGTAGAGTAGAAGACAGTA  
  
SerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsnAlaVal  
482 TCAAAGAAGAAGTGCACGAACTCGCCGCAAAGCTGGTCGCATTGGGCATCAATGCCGTG  
AGTTTCTTCTTCACGCTGCTTGAGCGGCGTTTCGACCAGCGTAACCCGTAGTTACGGCAC  
  
AlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValValValVal  
542 GCCTACTACCGGGTCTTGACGTGTCCGTATCCCGACCAGCGGCGATGTTGTGTCGTCGTG  
CGGATGATGGCGCCAGAACTGCACAGGCAGTAGGGCTGGTCGCCGCTACAACAGCAGCAC  
^ ^  
550 SAC2, 560 DRD1,  
  
AlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsn  
602 GCAACCGATGCCCTCATGACCGGCTATACCGGCGACTTCGACTCGGTGATAGACTGCAAT  
CGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTGACGTTA  
^  
615 BSPH1,  
  
ThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGluThrIle

FIG. 14A



662 ACGTGTGTCACCCAGACAGTCGATTTTCAGCCTTGACCCTACCTTCACCATTGAGACAATC  
TGCACACAGTGGGTCTGTCTAGCTAAAGTCGGAAGTGGGATGGAAGTGGTAACCTCTGTTAG

ThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLys  
722 ACGTCCCCCAAGATGCTGTCTCCCGCACTCAACGTCGGGGCAGGACTGGCAGGGGGAAG  
TGCGAGGGGGTTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCCTGACCGTCCCCCTTC

ProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSer  
782 CCAGGCATCTACAGATTTGTGGCACCAGGGGAGCGCCCCCTCCGGCATGTTGACTCGTCC  
GGTCCGTAGATGTCTAAACACCGTGCCCCCTCGCGGGAGGCCGTACAAGCTGAGCAGG

816 BGLI, 833 DRD1,

ValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThr  
842 GTCCTCTGTGAGTGCTATGACGCAGGCTGTGCTTGGTATGAGCTCACGCCCCGCGAGACT  
CAGGAGACACTCACGATACTGCGTCCGACACGAACCATACTCGAGTGCGGGCGGCTCTGA

881 SACI,

ThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeu  
902 ACAGTTAGGCTACGAGCGTACATGAACACCCCGGGGCTTCCCGTGTGCCAGGACCATCTT  
TGTCATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCACACGGTCTCTGGTAGAA

931 SMAI XMAI,

GluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGln  
962 GAATTTTGGGAGGGCGTCTTTACAGGCCTCACTCATATAGATGCCCACTTTCTATCCCAG  
CTTAAACCCCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGATAGGGTC

985 STUI,

ThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAla  
1022 ACAAAGCAGAGTGGGGAGAACCTTCTTACCTGGTAGCGTACCAAGCCACCGTGTGCGCT  
TGTTTCGTCTCACCCCTCTTGAAGGAATGGACCATCGCATGGTTCGGTGGCACACGCGA

1069 DRA3,

ArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysLeuIleArgLeuLys  
1082 AGGGCTCAAGCCCCCTCCCCATCGTGGGACCAGATGTGGAAGTGTGATTGCGCTCAAG  
TCCCGAGTTCGGGGAGGGGGTAGCACCTGGTCTACACCTTCACAACTAAGCGGAGTTC

ProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsnGluIle  
1142 CCCACCCTCCATGGGCCAACACCCCTGCTATACAGACTGGGCGCTGTTGAGAATGAAATC  
GGGTGGGAGGTACCCGTTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTACTTTAG

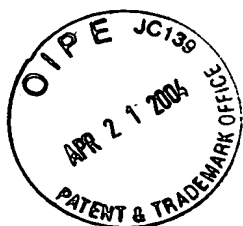
1150 NCOI,

ThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeuGluVal  
1202 ACCCTGACGCACCCAGTCACCAATACATCATGACATGCATGTGCGCCGACCTGGAGGTC  
TGGGACTGCGTGGGTGAGTGGTTTATGTAGTACTGTACGTACAGCCGGCTGGACCTCCAG

1230 BSPH1, 1234 DRD1, 1237 AVA3, 1245 EAG1 XMA3, 1250 DRD1,

ValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyrCysLeu  
1262 GTCACGAGCACCTGGGTGCTCGTTGGCGGCGTCCTGGCTGCTTTGGCCGCGTATTGCCTG  
CAGTGCTCGTGGACCCACGAGCAACCGCCGAGGACCGACGAAACCGCGCATAACGGAC

FIG. 14B



1322 S rThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAlaIleIle  
TCAACAGGCTGCGTGGTCATAGTGGGCAGGGTCGTCTTGTCGCGGAAGCCGGCAATCATA  
AGTTGTCCGACGCACCAGTATCACCCGTCCAGCAGAACAGGCCCTTCGGCCGTTAGTAT  
^

1369 NAEI,

1382 ProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGlnHisLeu  
CCTGACAGGGAAGTCCTCTACCGAGAGTTTCGATGAGATGGAAGAGTGCTCTCAGCACTTA  
GGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCACGAGAGTCGTGAAT  
^

1385 DRD1,

1442 ProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeuGlyLeu  
CCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGCAGAAGGCCCTCGGCCTC  
GGCATGTAGCTCGTTCCTACTACGAGCGGCTCGTCAAGTTCGTCTTCGGGAGCCGGAG

1502 LeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsnTrpGln  
CTGCAGACCGCGTCCCGTCAGGCAGAGGTTATCGCCCTGCTGTCCAGACCAACTGGCAA  
GACGTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTGACCGTT  
^ ^

1502 PSTI, 1507 TTH3I,

1562 LysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeu  
AAACTCGAGACCTTCTGGGCGAAGCATATGTGGAACCTTCATCAGTGGGATACAATACTTG  
TTTGAGCTCTGGAAGACCCGCTTCGTATACACCTTGAAGTAGTCACCCATGTATTGAAC  
^ ^

1565 XHOI, 1586 NDEI,

1622 AlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPheThrAla  
GCGGGCTTGTCAACGCTGCCTGGTAACCCCGCCATTGCTTCATTGATGGCTTTTACAGCT  
CGCCCGAACAGTTGCGACGGACCATTGGGGCGGTAACGAAGTAACCTACCGAAAATGTCGA  
^

1643 BSTE2, 1677 ALWN1 PVU2,

1682 AlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGlyGlyTrp  
GCTGTCAACAGCCCACTAACCCTAGCCAAACCTCCTCTTCAACATATTGGGGGGGTGG  
CGACAGTGGTCGGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCCCCACC

1742 ValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGlyLeuAla  
GTGGCTGCCCAGCTCGCCGCCCCCGGTGCCGCTACTGCCTTTGTGGGCGCTGGCTTAGCT  
CACCGACGGGTCGAGCGGCGGGGGCCACGGCGATGACGGAAACACCGCGACCGAATCGA  
^

1794 ESP1,

1802 GlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyr  
GGCGCCGCCATCGGCAGTGTGGACTGGGGAAGGTCCTCATAGACATCCTTGAGGGTAT  
CCGCGGCGGTAGCCGTCACAACCTGACCCCTTCCAGGAGTATCTGTAGGAACGTCCCATA  
^

1802 KAS1 NARI,

1862 GlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluValProSer  
GGCGCGGGCGTGGCGGAGCTCTTGTTGGCATTCAAGATCATGAGCGGTGAGGTCCCTCC  
CCGCGCCCGCACCGCCCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAGGGGAGG  
^ ^

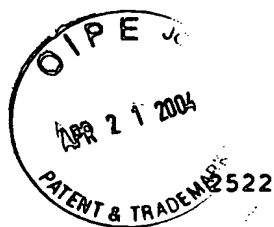
1878 SACI, 1899 BSPH1,

FIG. 14C



1922 ThrGluAspLeuValAsnLeuLeuProAlaIle LeuSerProGlyAlaLeuValValGly  
ACGGAGGACCTGGTCAATCTACTGCCCGCCATCCTCTCGCCCGGAGCCCTCGTAGTCGGC  
TGCCTCCTGGACCAGTTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCATCAGCCG  
1928 TTH3I,  
1982 ValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrp  
GTGGTCTGTGCAGCAATACTGCGCCGGCACGTTGGCCCGGGCGAGGGGGCAGTGCAGTGG  
CACCAGACACGTCGTTATGACGCGGCCGTGCAACCGGGCCCGCTCCCCCGTCACGTCACC  
2004 NAEI, 2017 SMAI XMAI,  
2042 MetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrVal  
ATGAACCGGCTGATAGCCTTCGCCTCCCGGGGAACCATGTTTCCCCCAGCACTACGTG  
TACTTGGCCGACTATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGGTGC GTGATGCAC  
2067 SMAI XMAI, 2093 DRA3,  
2102 ProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGln  
CCGGAGAGCGATGCAGCTGCCCCGCGTCACTGCCATACTCAGCAGCCTCACTGTAACCCAG  
GGCCTCTCGCTACGTGACGGGCGCAGTGACGGTATGAGTCGTGCGAGTGACATTGGGTC  
2115 PVU2, 2159 ALWN1,  
2162 LeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSerGlySer  
CTCCTGAGGCGACTGCACCAGTGGATAAGCTCGGAGTGTACCACTCCATGCTCCGGTTCC  
GAGGACTCCGCTGACGTGGTCACCTATTTCGAGCCTCACATGGTGAGGTACGAGGCCAAGG  
2164 MST2, 2220 ECON1,  
2222 TrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeu  
TGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCGACTTTAAGACCTGGCTA  
ACCGATTCCCTGTAGACCCTGACCTATACGCTCCACAACCTCGCTGAAATTCTGGACCGAT  
2282 LysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyr  
AAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCCTTTGTGTCCTGCCAGCGCGGGTAT  
TTTCGATTTCGAGTACGGTGTGACGGACCCTAGGGGAAACACAGGACGGTCGCGCCCAT  
2285 ESP1, 2300 PVU2, 2310 BAMHI,  
2342 LysGlyValTrpArgGlyAspGlyIleMetHisThrArgCysHisCysGlyAlaGluIle  
AAGGGGGTCTGGCGAGGGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCTGAGATC  
TTCCCCCAGACCGCTCCCCCTGCCGTAGTACGTGTGAGCGACGGTGACACCTCGACTCTAG  
2402 ThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArgAsnMet  
ACTGGACATGTCAAAAACGGGACGATGAGGATCGTCGGTCTAGGACCTGCAGGAACATG  
TGACCTGTACAGTTTTTGCCCTGCTACTCCTAGCAGCCAGGATCCTGGACGTCCTTGAC  
2425 BSAB1, 2441 AVR2, 2448 SSE83871, 2449 PSTI,  
2462 TrpSerGlyThrPheProIleAsnAlaTyrThrThrGlyProCysThrProLeuProAla  
TGGAGTGGGACCTTCCCCATTAATGCCTACACCACGGGCCCCCTGTACCCCCCTTCTGCG  
ACCTCACCTGGAAGGGGTAATTACGGATGTGGTGCCCGGGGACATGGGGGGAAGGACGC  
2480 ASE1, 2497 APAI,  
ProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIleArgGln

FIG. 14D



2522 CCGAACTACACGTTTCGCGCTATGGAGGGTGTCTGCAGAGGAATACGTGGAGATAAGGCAG  
GGCTTGATGTGCAAGCGCGATACCTCCACAGACGTCTCCTTATGCACCTCTATTCCGTC  
2553 PSTI,  
ValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysProCysGln  
2582 GTGGGGGACTTCCACTACGTGACGGGTATGACTACTGACAATCTTAAATGCCCGTGCCAG  
CACCCCCTGAAGGTGATGCACTGCCATACTGATGACTGTTAGAATTTACGGGCACGGTC  
2594 DRA3,  
ValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPheAlaPro  
2642 GTCCCATCGCCGAATTTTTACAGAATTGGACGGGGTGCGCCTACATAGGTTTGCGCCC  
CAGGGTAGCGGGCTTAAAAAGTGTCTTAACCTGCCCCACGCGGATGTATCCAAACGCGGG  
ProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGluTyrPro  
2702 CCCTGCAAGCCCTTGCTGCGGGAGGAGGTATCATTGAGAGTAGGACTCCACGAATACCCG  
GGGACGTTGCGGAACGACGCCCTCTCCATAGTAAGTCTCATCCTGAGGTGCTTATGGGC  
2757 HGIE2,  
ValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSerMetLeu  
2762 GTAGGGTCGCAATTACCTTGCGAGCCCGAACCGGACGTGGCCGTGTTGACGTCCATGCTC  
CATCCCAGCGTTAATGGAACGCTCGGGCTTGGCCTGCACCGGCACAACCTGCAGGTACGAG  
2809 AAT2,  
ThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGlySerPro  
2822 ACTGATCCCTCCCATATAACAGCAGAGGCGGCCGGCGAAGGTTGGCGAGGGGATCACCC  
TGACTAGGGAGGGTATATTGTCGTCTCCGCCGGCCCGCTTCCAACCGCTCCCCTAGTGGG  
2850 EAG1 XMA3,  
ProSerValAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuLysAlaThrCys  
2882 CCCTCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCTCAAGGCAACTTGC  
GGGAGACACCGGTCGAGGAGCCGATCGGTGATAGGCGAGGTAGAGAGTTCCGTTGAACG  
2889 BALI, 2903 NHEI,  
ThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrpArgGln  
2942 ACCGCTAACCATGACTCCCCTGATGCTGAGCTCATAGAGGCCAACCTCCTATGGAGGCAG  
TGGCGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGGTTGGAGGATACCTCCGTC  
2966 ESP1, 2969 SACI,  
GluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeuAspSer  
3002 GAGATGGGCGGCAACATCACCAGGGTTGAGTCAGAAAACAAAGTGGTGATTCTGGACTCC  
CTCTACCCGCCGTTGTAGTGGTCCCAACTCAGTCTTTTGTTCACCACTAAGACCTGAGG  
PheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGluIleLeu  
3062 TTCGATCCGCTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGCAGAAATCCTG  
AAGCTAGGCGAACACCGCCTCCTCTGCTCGCCCTCTAGAGGCATGGGCGTCTTTAGGAC  
3096 BGL2,  
ArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyrAsnPro  
3122 CGGAAGTCTCGGAGATTCGCCAGGCCCTGCCCGTTTGGGCGCGGCCGACTATAACCCC

FIG. 14E



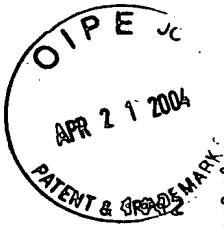
GCCTTCAGAGCCTCTAAGCGGGTCCGGGACGGGCAAACCCGCGCCGGCCTGATATTGGGG  
3143 ALWN1, 3164 EAG1 XMA3,  
ProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGlyCysPro  
3182 CCGCTAGTGGAGACGTGGAAAAAGCCCGACTACGAACCACCTGTGGTCCATGGCTGCCCCG  
GGCGATCACCTCTGCACCTTTTTCGGGCTGATGCTTGGTGGACACCAGGTACCGACGGGC  
3217 HGIE2, 3229 NCOI,  
LeuProProProLysSer roValProProProArgLysLysArgThrValValLeu  
3242 CTTCACCTCCAAAGTCCCCCTCTGTGCCTCCGCCTCGGAAGAAGCGGACGGTGGTCCTC  
GAAGGTGGAGGTTTCAGGGGAGGACACGGAGGCGGAGCCTTCTTCGCCTGCCACCAGGAG  
ThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGlySerSer  
3302 ACTGAATCAACCCTATCTACTGCCTTGGCCGAGCTCGCCACCAGAAGCTTTGGCAGCTCC  
TGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGGTCTTCGAAACCGTCGAGG  
3332 SACI, 3346 HIND3,  
SerThrSerGlyIleThrGlyAspAsnThrThrThrSerSerGluProAlaProSerGly  
3362 TCAACTTCCGGCATTACGGGCGACAATACGACAACATCCTCTGAGCCCGCCCTTCTGGC  
AGTTGAAGGCCGTAATGCCCGCTGTTATGCTGTTGTAGGAGACTCGGGCGGGGAAGACCG  
CysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGlyGluPro  
3422 TGCCCCCCCCGACTCCGACGCTGAGTCCTATTCTCCATGCCCCCCTGGAGGGGGAGCCT  
ACGGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGGACCTCCCCCTCGGA  
3437 EAM11051,  
GlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsnAlaGlu  
3482 GGGGATCCGGATCTTAGCGACGGGTCATGGTCAACGGTCAGTAGTGAGGCCAACGCGGAG  
CCCCTAGGCCTAGAATCGCTGCCAGTACCAGTTGCCAGTCATCACTCCGGTTGCGCCTC  
3484 BAMHI, 3485 BSAB1, 3487 BSPE1,  
AspValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuValThrProCysAla  
3542 GATGTCGTGTGCTGCTCAATGTCTTACTCTTGGACAGGCGCACTCGTCACCCCGTGCGCC  
CTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCGCGTGAGCAGTGGGGCACGCGG  
3589 DRA3, 3600 SAC2,  
AlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsn  
3602 GCGGAAGAACAGAACTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTCACCACAAT  
CGCCTTCTTGTCTTTGACGGGTAGTTACGTGATTCGTTGAGCAACGATGCAGTGGTGTTA  
3611 ALWN1, 3655 PFLM1,  
LeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAsp  
3662 TTGGTGTATTCCACCACCTCACGCAGTGCTTGCCAAAGGCAGAAGAAAGTCACATTTGAC  
AACCACATAAGGTGGTGGAGTGCGTCACGAACGGTTTCCGTCTTCTTTCAGTGTAAGT  
3681 DRA3,  
ArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAlaAlaAla  
3722 AGACTGCAAGTTCTGGACAGCCATTACCAGGACGTAAGGAGGTTAAAGCAGCGGCG  
TCTGACGTTCAAGACCTGTGCGTAATGGTCTGATGAGTTCTCCAATTTCTGTCGCGCG

FIG. 14F



3782 SerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysS rLeuThrProProHis  
TCAAAAGTGAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGCAGCCTGACGCCCCACAC  
AGTTTTCACTCCGATTGAACGATAGGCATCTCCTTCGAACGTCGGACTGCGGGGGTGTG  
3816 HIND3,  
3842 SerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAla  
TCAGCCAAATCCAAGTTTGGTTATGGGGCAAAGACGTCCGTTGCCATGCCAGAAAGGCC  
AGTCGGTTTAGGTTCAAACCAATACCCCGTTTTCTGCAGGCAACGGTACGGTCTTTCCGG  
3875 AAT2, 3890 BGLI,  
3902 ValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrProIleAsp  
GTAACCCACATCAACTCCGTGTGGAAAGACCTTCTGGAAGACAATGTAACACCAATAGAC  
CATTTGGGTGTAGTTGAGGCACACCTTTCTGGAAGACCTTCTGTTACATTGTGGTTATCTG  
3962 ThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGlyArgLys  
ACTACCATCATGGCTAAGAACGAGGTTTTCTGCGTTCAGCCTGAGAAGGGGGTTCGTAAG  
TGATGGTAGTACCGATTCTTGCTCCAAAGACGCAAGTCGGACTCTTCCCCCAGCATTCT  
4022 ProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMetAlaLeu  
CCAGCTCGTCTCATCGTGTTCCTCCCGATCTGGGCGTGCGCGTGTGCGAAAAGATGGCTTTG  
GGTCGAGCAGAGTAGCACAAAGGGGCTAGACCCGCACGCGCACAGCTTTTCTACCGAAAC  
4082 TyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPheGlnTyr  
TACGACGTGGTTACAAAGCTCCCCTTGGCCGTGATGGGAAGCTCCTACGGATTCCAATAC  
ATGCTGCACCAATGTTTCGAGGGGAACCGGCACTACCCTTCGAGGATGCCTAAGGTTATG  
4142 SerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThrProMet  
TCACCAGGACAGCGGGTTGAATTCCTCGTGCAAGCGTGGAAGTCCAAGAAAACCCCAATG  
AGTGGTCTCTGCGCCCACTTAAGGAGCACGTTTCGCACCTTCAGGTTCTTTTGGGGTTAC  
4160 ECORI,  
4202 GlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIleArgThr  
GGGTTCCTCGTATGATACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATCCGTACG  
CCCAAGAGCATACTATGGGCGACGAACTGAGGTGTCAGTGACTCTCGCTGTAGGCATGC  
4229 DRD1, 4236 ALWN1,  
4262 GluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIleLysSer  
GAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCCAAGCCCGGTGGCCATCAAGTCC  
CTCCTCCGTTAGATGGTTACAACACTGGAGCTGGGGGTTTCGGGCGCACCGGTAGTTCAGG  
4301 BGLI, 4308 BALI,  
4322 LeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsnCysGly  
CTCACCAGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAGAAGTGCAGG  
GAGTGGCTCTCCGAAATACAACCCCGGGAGAATGGTTAAGTTCCCCCTCTTGACGCCG  
4345 APAI,  
4382 TyrArgArgCysArgAlaSerGlyValLeuThrThrSerCysGlyAsnThrLeuThrCys  
TATCGCAGGTGCCGCGCGAGCGGCGTACTGACAACTAGCTGTGGTAACACCCTCACTTGC  
ATAGCGTCCACGGCGCGCTCGCCGCATGACTGTTGATCGACACCATTGTGGGAGTGAACG

FIG. 14G



TyrIleLysAlaArgAlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMetLeuVal  
TACATCAAGGCCCGGGCAGCCTGTCTGAGCCGAGGGCTCCAGGACTGCACCATGCTCGTG  
ATGTAGTTCCGGGCGCCGTCGGACAGCTCGGCGTCCCGAGGTCCTGACGTGGTACGAGCAC  
^  
4452 SMAI XMAI,  
  
CysGlyAspAspLeuValValIleCysGluSerAlaGlyValGlnGluAspAlaAlaSer  
4502 TGTGGCGACGACTTAGTCGTTATCTGTGAAAGCGCGGGGTCCAGGAGGACGCGGCGAGC  
ACACCGCTGCTGAATCAGCAATAGACACTTTCGCGCCCCCAGGTCCTCCTGCGCCGCTCG  
^ ^  
4508 DRD1, 4511 TTH3I,  
  
LeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProGln  
4562 CTGAGAGCCTTCACGGAGGCTATGACCAGGTACTCCGCCCCCCTGGGGACCCCCACAA  
GACTCTCGGAAGTGCCCTCCGATACTGGTCCATGAGGCGGGGGGACCCCTGGGGGGTGTT  
  
ProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaHisAsp  
4622 CCAGAATACGACTTGGAGCTCATAACATCATGCTCCTCCAACGTGTGAGTCGCCCACGAC  
GGTCTTATGCTGAACCTCGAGTATTGTAGTACGAGGAGGTTGCACAGTCAGCGGGTGCTG  
^  
4637 SACI,  
  
GlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAlaArgAla  
4682 GCGCTGGAAAGAGGGTCTACTACCTACCCGTGACCCTACAACCCCTCGCGAGAGCT  
CCGCGACCTTCTCCAGATGATGGAGTGGGCACTGGGATGTTGGGGGGAGCGCTCTCGA  
^  
4731 NRUI,  
  
AlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIleMetPhe  
4742 GCGTGGGAGACAGCAAGACACACTCCAGTCAATTCCTGGCTAGGCAACATAATCATGTTT  
CGCACCTCTGTCTTCTGTGTGAGGTGAGTTAAGGACCGATCCGTTGTATTAGTACAAA  
  
AlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePheSerValLeuIleAla  
4802 GCCCCACACTGTGGGCGAGGATGATACTGATGACCCATTTCTTTAGCGTCCTTATAGCC  
CGGGGCTGTGACACCCGCTCCTACTATGACTACTGGGTAAAGAAATCGCAGGAATATCGG  
^^  
4806 PFLM1, 4807 DRA3,  
  
ArgAspGlnLeuGluGlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSerIleGlu  
4862 AGGGACAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGGCCTGCTACTCCATAGAA  
TCCCTGGTTCGAACCTGTCCGGGAGCTAACGCTCTAGATGCCCCGACGATGAGGTATCTT  
^  
4893 BGL2,  
  
ProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSerLeuHis  
4922 CCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTTCACTCCAC  
GGTGACCTAGATGGAGGTTAGTAAGTTTCTGAGGTACCGGAGTCGCGTAAAGTGAAGTG  
^  
4954 NCOI,  
  
SerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGlyValPro  
4982 AGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGCATGCCTCAGAAAACCTGGGGTACCG  
TCAATGAGAGGTCCACTTTAGTTATCCACCGCGTACGGAGTCTTTTGACCCCATGGC  
^  
5015 SPHI, 5035 KPNI,  
  
ProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAlaArgGly

FIG. 14H



5042 CCCTTGCGAGCTTGGAGACACCGGGCCCGGAGCGTCCGCGCTAGGCTTCTGGCCAGAGGA  
GGGAACGCTCGAACCTCTGTGGCCCGGGCCTCGCAGGCGCGATCCGAAGACCGGTCTCCT  
5064 APAI, 5091 BALI,  
GlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLys  
5102 GGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAAGCTCAAA  
CCGTCCCACGGTATACACCGTTCATGGAGAAGTTGACCCGTCATTCTTGTTTCGAGTTT  
5113 NDEI,  
LeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAlaGlyTyr  
5162 CTCACTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCCGGCTGGTTCACGGCTGGCTAC  
GAGTGAGGTTATCGCCGGCGACCGGTTCGACCTGAACAGGCCGACCAAGTGCCGACCGATG  
5174 NOTI, 5175 EAG1 XMA3, 5182 BALI, 5186 PVU2,  
SerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCys  
5222 AGCGGGGGAGACATTTATCACAGCGTGTCTCATGCCCGGCCCCGCTGGATCTGGTTTTGC  
TCGCCCCCTCTGTAAATAGTGTGCGACAGAGTACGGGCCGGGGCGACCTAGACCAAACG  
5240 DRA3,  
LeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgOP  
5282 CTACTCCTGCTTGCTGCAGGGGTAGGCATCTACCTCCTCCCCAACCGATGAATAGTCGAC  
GATGAGGACGAACGACGTCCCCATCCGTAGATGGAGGAGGGGTGGCTACTTATCAGCTG  
5295 PSTI, 5336 SALI,

FIG. 14I

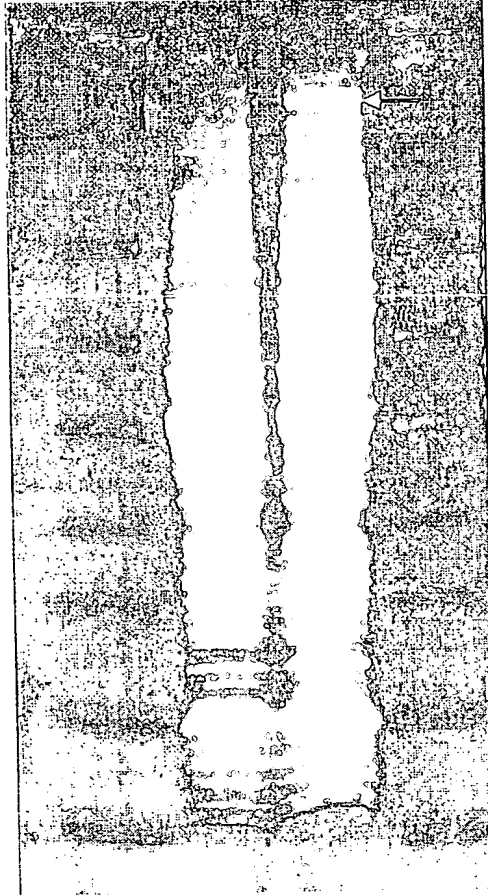


FIG. 15

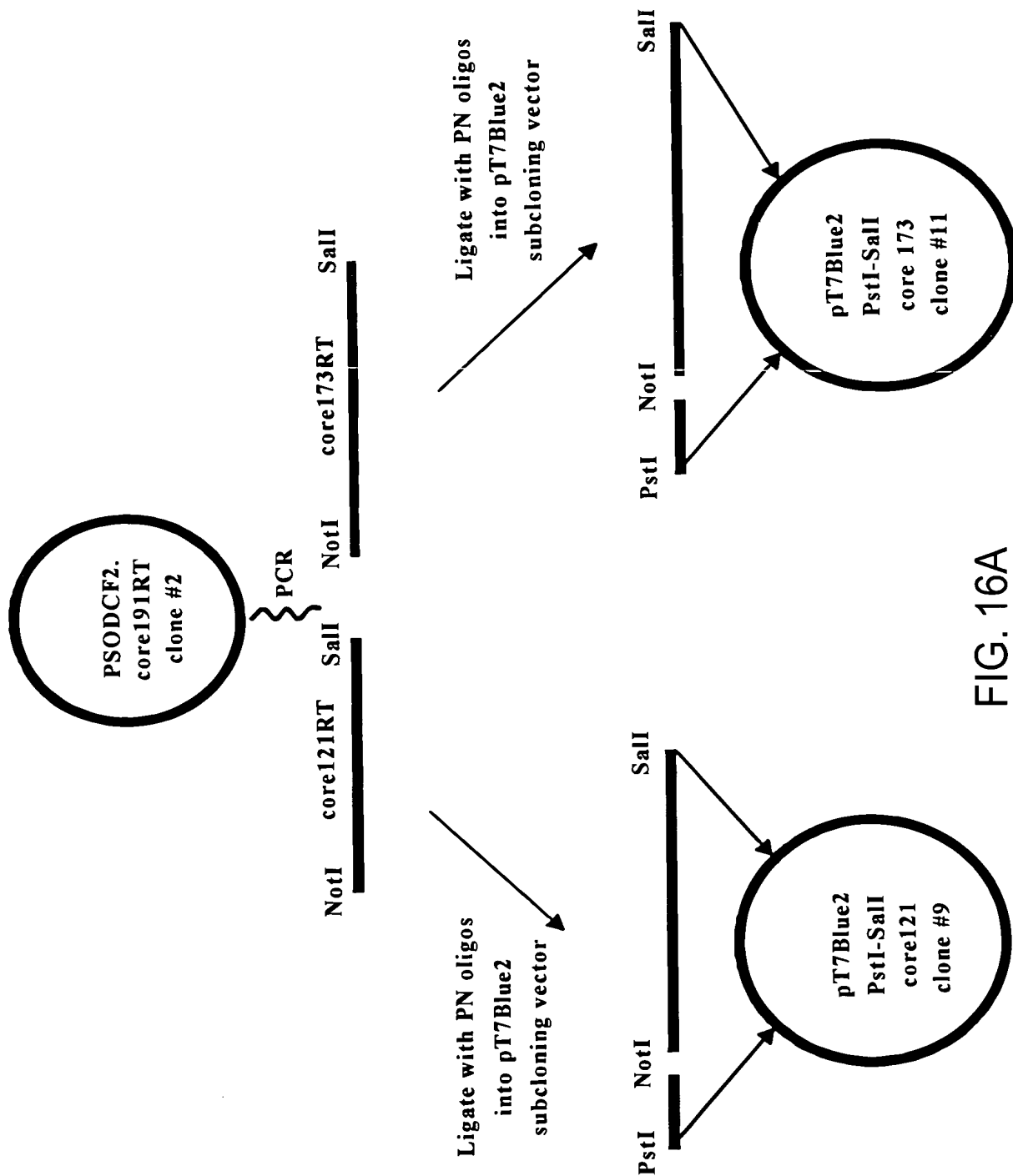


FIG. 16A

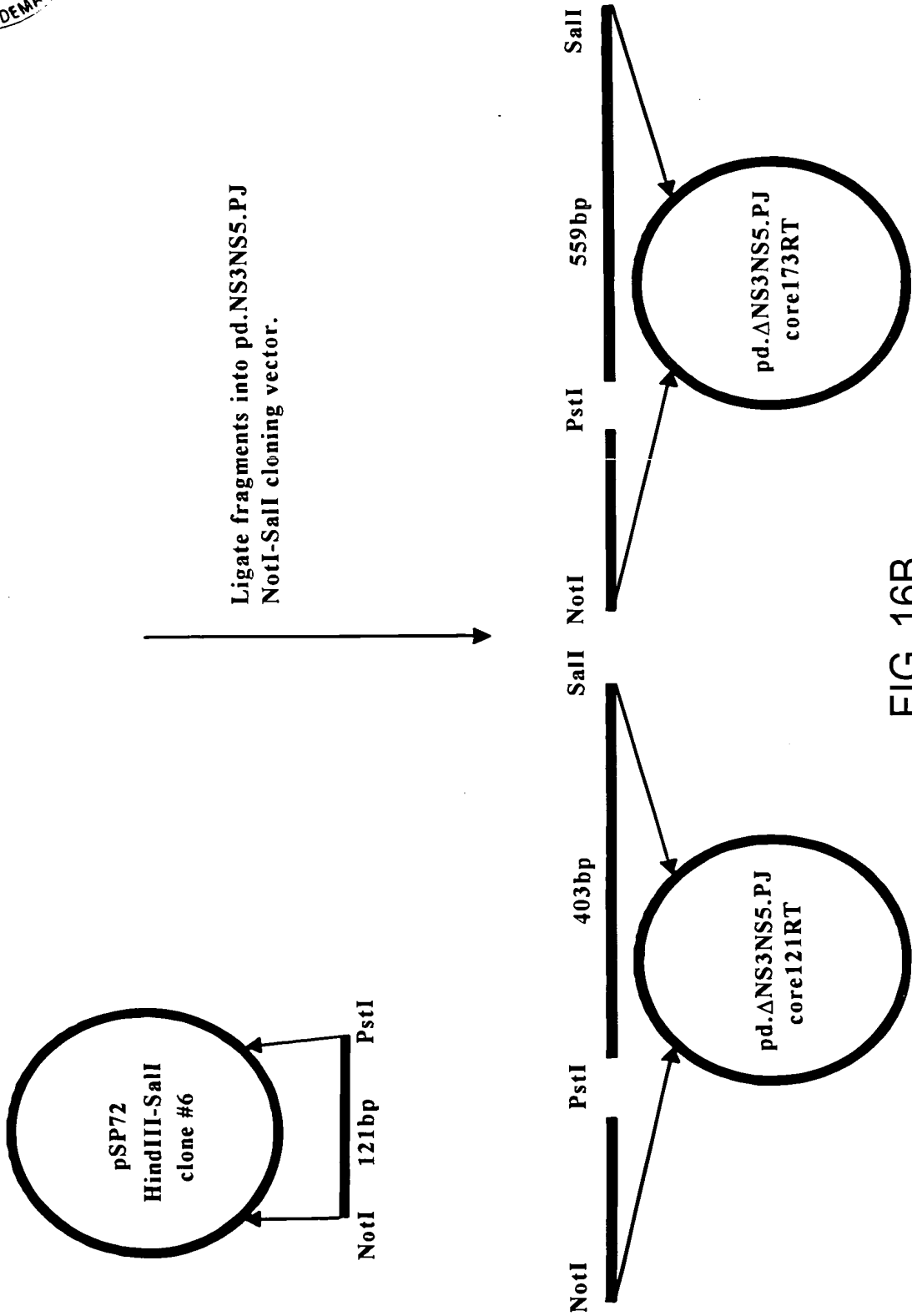


FIG. 16B



MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsn  
2 AGCTTACAAAACAAAATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTACTCAAC  
TCGAATGTTTTGTTTTACCGACGTATACGTCGAGTCCCGATATTCCACGATCATGAGTTG  
^ ^ ^  
1 HIND3, 24 NDEI, 52 SCAI,  
  
ProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGlyIleAsp  
62 CCCTCTGTTGCTGCAACACTGGGCTTTGGTGCTTACATGTCCAAGGCTCATGGGATCGAT  
GGGAGACAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCCTAGCTA  
^  
116 CLAI,  
  
ProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyrSerThr  
122 CCTAACATCAGGACCGGGTGAGAACAATTACCACTGGCAGCCCCATCACGTACTCCACC  
GGATTGTAGTCCTGGCCCCACTCTTGTTAATGGTGACCGTCGGGGTAGTGCATGAGGTGG  
  
TyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIleIleCys  
182 TACGGCAAGTTCCTTGCCGACGGCGGGTGCTCGGGGGGCGCTTATGACATAATAATTTGT  
ATGCCGTTCAAGGAACGGCTGCCGCCCCACGAGCCCCCGGAATACTGTATTATTAAACA  
  
AspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGln  
242 GACGAGTGCCACTCCACGGATGCCACATCCATCTTGGGCATTGGCACTGTCCTTGACCAA  
CTGCTCACGGTGAGGTGCCTACGGTGTAGGTAGAACCCGTAACCGTGACAGGAAGTGGTT  
  
AlaGluThrAlaGlyAlaArgLeuValValLeuAlaThrAlaThrProProGlySerVal  
302 GCAGAGACTGCGGGGGCGAGACTGGTTGTGCTCGCCACCGCCACCCCTCCGGGCTCCGTC  
CGTCTCTGACGCCCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCCGAGGCAG  
^  
303 ALWN1,  
  
ThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIleProPhe  
362 ACTGTGCCCCATCCCAACATCGAGGAGGTTGCTCTGTCCACCACCGGAGAGATCCCTTTT  
TGACACGGGGTAGGGTTGTAGCTCCTCCAACGAGACAGGTGGTGGCCTCTCTAGGGAAAA  
  
TyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePheCysHis  
422 TACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGGAGACATCTCATCTTCTGTCAT  
ATGCCGTTCCGATAGGGGGAGCTTCATTAGTTCCCCCCTCTGTAGAGTAGAAGACAGTA

FIG. 17A



482 SerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsnAlaVal  
TCAAAGAAGAAGTGGCAGCACTCGCCGCAAAGCTGGTCGCATTGGGCATCAATGCCGTG  
AGTTTCTTCTTCACGCTGCTTGAGCGGCGTTTCGACCAGCGTAACCCGTAGTTACGGCAC

542 AlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValValValVal  
GCCTACTACCGCGGTCTTGACGTGTCCGTATCCCGACCAGCGGCGATGTTGTCTCGTCTG  
CGGATGATGGCGCCAGAACTGCACAGGCAGTAGGGCTGGTCGCCGCTACAACAGCAGCAC  
550 SAC2, 560 DRD1,

602 AlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsn  
GCAACCGATGCCCTCATGACCGGCTATACCGGCGACTTCGACTCGGTGATAGACTGCAAT  
CGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTGACGTTA  
615 BSPH1,

662 ThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGluThrIle  
ACGTGTGTCAACCCAGACAGTCGATTTTCAGCCTTGACCTACCTTCACCATTGAGACAATC  
TGCACACAGTGGGTCTGTCTAGCTAAAGTCGGAAGTGGGATGGAAGTGGTAAGTCTGTTAG

722 ThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLys  
ACGCTCCCCAAGATGCTGTCTCCCGCACTCAACGTCGGGGCAGGACTGGCAGGGGGGAG  
TGCGAGGGGGTTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCTGACCGTCCCCCTTC

782 ProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSer  
CCAGGCATCTACAGATTTGTGGCACCGGGGGAGCGCCCCCTCCGGCATGTTTCGACTCGTCC  
GGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAGCTGAGCAGG  
816 BGLI, 833 DRD1,

842 ValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThr  
GTCCTCTGTGAGTGCTATGACGCAGGCTGTGCTTGGTATGAGCTCACGCCCGCCGAGACT  
CAGGAGACACTCACGATACTGCGTCCGACACGAACCATACTCGAGTGCGGGCGGCTCTGA  
881 SACI,

902 ThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeu  
ACAGTTAGGCTACGAGCGTACATGAACACCCCGGGGCTTCCCGTGTGCCAGGACCATCTT  
TGTCATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCACACGGTCTGTTAGAA  
931 SMAI XMAI,

962 GluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGln  
GAATTTTGGGAGGGCGTCTTTACAGGCCTCACTCATATAGATGCCACTTTCTATCCCAG  
CTTAAACCCCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGATAGGGTC  
985 STUI,

1022 ThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAla  
ACAAAGCAGAGTGGGGAGAACCTTCTTACCTGGTAGCGTACCAAGCCACCGTGTGCGCT  
TGTTTCGTCTCACCCCTCTTGAAGGAATGGACCATCGCATGGTTTCGGTGGCACACGCGA  
1069 DRA3,

1082 ArgAlaGlnAlaProPr ProS rTrpAspGlnMetTrpLysCysLeuIleArgLeuLys  
AGGGCTCAAGCCCCCTCCCCCATCGTGGGACCAGATGTGGAAGTGTGTTGATTCGCCTCAAG

FIG. 17B



TCCCGAGTTCGGGGAGGGGGTAGCACCTGGTCTACACCTTCACAACTAAGCGGAGTTC

ProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsnGluIle  
 1202 CCCACCTCCATGGGCCAACACCCCTGCTATACAGACTGGGCGCTGTTTCAAGTAAATC  
 GGGTGGGAGGTACCCGGTTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTACTTTAG  
 1150 NCOI,

ThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeuGluVal  
 1202 ACCCTGACGCACCCAGTCACCAAATACATCATGACATGCATGTCGGCCGACCTGGAGGTC  
 TGGGACTGCGTGGGTGAGTGGTTTATGTAGTACTGTACGTACAGCCGGCTGGACCTCCAG  
 1230 BSPH1, 1234 DRD1, 1237 AVA3, 1245 EAG1 XMA3, 1250 DRD1,

ValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyrCysLeu  
 1262 GTCACGAGCACCTGGGTGCTCGTTGGCGGGCTCCTGGCTGCTTTGGCCGCGTATTGCCTG  
 CAGTGCTCGTGGACCCACGAGCAACCGCCGAGGACCGACGAAACCGGCGCATAACGGAC

SerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAlaIleIle  
 1322 TCAACAGGCTGCGTGGTCATAGTGGGCAGGGTCGTCTTGTCCGGGAAGCCGGCAATCATA  
 AGTTGTCCGACGCACCAGTATCACCCGTCCCAGCAGAACAGGCCCTTCGGCCGTTAGTAT  
 1369. NAEI,

ProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGlnHisLeu  
 1382 CCTGACAGGGAAGTCCTCTACCGAGAGTTCGATGAGATGGAAGAGTGCTCTCAGCACTTA  
 GGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCACGAGAGTCGTGAAT  
 1385 DRD1,

ProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeuGlyLeu  
 1442 CCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAGCAGAAGGCCCTCGGCCTC  
 GGCATGTAGCTCGTTCCTACTACGAGCGGCTCGTCAAGTTCGTCTTCGGGAGCCGGAG

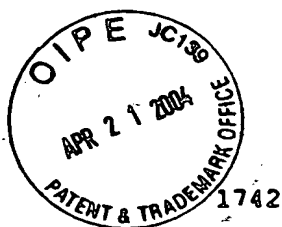
LeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsnTrpGln  
 1502 CTGCAGACCGCGTCCCGTCAGGCAGAGGTTATCGCCCTGCTGTCCAGACCAACTGGCAA  
 GACGTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTGACCGTT  
 1502 PSTI, 1507 TTH3I,

LysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeu  
 1562 AAACCTCGAGACCTTCTGGGCGAAGCATATGTGGAACCTCATCAGTGGGATACAATACTTG  
 TTTGAGCTCTGGAAGACCCGCTTCGTATACACCTTGAAGTAGTCACCCTATGTTATGAAC  
 1565 XHOI, 1586 NDEI,

AlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPheThrAla  
 1622 GCGGGCTTGTCAACGCTGCCTGGTAACCCCGCCATTGCTTCATTGATGGCTTTTACAGCT  
 CGCCCGAACAGTTGCGACGGACCAATTGGGGCGGTAACGAAGTAACTACCGAAAATGTCGA  
 1643 BSTE2, 1677 ALWN1 PVU2,

AlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGlyGlyTrp  
 1682 GCTGTCAACGACCCACTAACCCTAGCCAAACCCTCCTCTTCAACATATTGGGGGGGTGG  
 CGACAGTGGTGGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCCCCACC

FIG. 17C



ValAlaAlaGlnLeuAlaAlaPr GlyAlaAlaThrAlaPheValGlyAlaGlyLeuAla  
1742 GTGGCTGCCAGCTCGCCGCCCCCGGTGCCGCTACTGCCCTTGTGGGCGCTGGCTTAGCT  
CACCGACGGGTGAGCGGGCGGGGCCACGGCGATGACGGAACACCCGCGACCGAATCGA  
1794 ESP1,  
GlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyr  
1802 GCGCGCCCATCGGCAGTGTGGACTGGGGAAGGTCTCATAGACATCCTTGAGGGTAT  
CCGCGGCGGTAGCCGTCACAACCTGACCCCTTCCAGGAGTATCTGTAGGAACGTCCATA  
1802 KAS1 NARI,  
GlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluValProSer  
1862 GCGCGGGCGTGGCGGGAGCTCTTGTGGCATTCAAGATCATGAGCGGTGAGGTCCCCTCC  
CCGCGCCCGCACCGCCCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAGGGGAGG  
1878 SACI, 1899 BSPH1,  
ThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuValValGly  
1922 ACGGAGGACCTGGTCAATCTACTGCCCGCCATCCTCTCGCCCGGAGCCCTCGTAGTCGGC  
TGCCTCCTGGACCAGTTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCATCAGCCG  
1928 TTH3I,  
ValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrp  
1982 GTGGTCTGTGCAGCAATACTGCGCCGGCACGTTGGCCCGGGCGAGGGGGCAGTGCAGTGG  
CACCAGACACGTCGTTATGACGCGGCCGTGCAACCGGGCCCGCTCCCCCGTCACGTCACC  
2004 NAEI, 2017 SMAI XMAI,  
MetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrVal  
2042 ATGAACCGGCTGATAGCCTTCGCCTCCCGGGGAACCATGTTTCCCCACGCACTACGTG  
TACTTGGCCGACTATCGGAAGCGGAGGCCCCCTTGGTACAAAGGGGGTGCGTGATGCAC  
2067 SMAI XMAI, 2093 DRA3,  
ProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGln  
2102 CCGGAGAGCGATGCAGCTGCCCGCGTCACTGCCATACTCAGCAGCCTCACTGTAACCCAG  
GGCCTCTCGCTACGTGACGGGCGCAGTGACGGTATGAGTCGTCGGAGTGACATTGGGTC  
2115 PVU2, 2159 ALWN1,  
LeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSerGlySer  
2162 CTCCTGAGGCGACTGCACCAGTGGATAAGCTCGGAGTGTACCACTCCATGCTCCGGTTCC  
GAGGACTCCGCTGACGTGGTCACCTATTCGAGCCTCACATGGTGAGGTACGAGGCCAAGG  
2164 MST2, 2220 ECON1,  
TrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeu  
2222 TGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCGACTTTAAGACCTGGCTA  
ACCGATTCCCTGTAGACCCTGACCTATACGCTCCACAACCTCGCTGAAATTCTGGACCGAT  
2282 LysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyr  
AAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCTTTGTGTCCTGCCAGCGCGGGTAT  
TTTCGATTGAGTACGGTGTGACGGACCCTAGGGGAACACAGGACGGTTCGCGCCCAT  
2285 ESP1, 2300 PVU2, 2310 BAMHI,

FIG. 17D



2342 LysGlyValTrpArgGlyAspGlyIleMetHisThrArgCysHisCysGlyAlaGluIle  
AAGGGGGTCTGGCGAGGGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCTGAGATC  
TTCCCCAGACCGCTCCCCTGCCGTAGTACGTGTGAGCGACGGTGACACCTCGACTCTAG

2402 ThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArgAsnMet  
ACTGGACATGTCAAAAACGGGACGATGAGGATCGTCGGTCTTAGGACCTGCAGGAACATG  
TGACCTGTACAGTTTTTGGCCTGCTACTCCTAGCAGCCAGGATCCTGGACGTCCTTGATC  
2425 BSAB1, 2441 AVR2, 2448 SSE83871, 2449 PSTI,

2462 TrpSerGlyThrPheProIleAsnAlaTyrThrThrGlyProCysThrProLeuProAla  
TGGAGTGGGACCTTCCCCATTAATGCCTACACCACGGGCCCTGTACCCCCCTTCTGCG  
ACCTCACCTGGAAGGGGTAAATTACGGATGTGGTGCCCGGGACATGGGGGAAGGACGC  
2480 ASE1, 2497 APAI,

2522 ProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIleArgGln  
CCGAACACACGTTTCGCGCTATGGAGGGTGTCTGCAGAGGAATACGTGGAGATAAGGCAG  
GGCTTGATGTGCAAGCGCGATACCTCCCACAGACGTCCTTATGCACCTCTATTCCGTC  
2553 PSTI,

2582 ValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysProCysGln  
GTGGGGGACTTCCACTACGTGACGGGTATGACTACTGACAATCTTAAATGCCCGTGCCAG  
CACCCCTGAAGGTGATGCACTGCCCATACTGATGACTGTTAGAATTTACGGGCACGGTC  
2594 DRA3,

2642 ValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPheAlaPro  
GTCCCATCGCCGAATTTTTACAGAATTGGACGGGGTGCGCCTACATAGGTTTGCGCCC  
CAGGGTAGCGGGCTTAAAAAGTGTCTTAACCTGCCCCACGGGATGTATCCAAACGCGGG

2702 ProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGluTyrPro  
CCCTGCAAGCCCTTGCTGCGGGAGGAGGTATCATTCAGAGTAGGACTCCACGAATACCCG  
GGGACGTTGCGGAACGACGCCCTCCTCCATAGTAAGTCTCATCCTGAGGTGCTTATGGGC  
2757 HGIE2,

2762 ValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSerMetLeu  
GTAGGGTCGCAATTACCTTGCGAGCCCGAACCGGACGTGGCCGTGTTGACGTCCATGCTC  
CATCCCAGCGTTAATGGAACGCTCGGGCTTGGCCTGCACCGGCACAACTGCAGGTACGAG  
2809 AAT2,

2822 ThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGlySerPro  
ACTGATCCCTCCCATATAACAGCAGAGGCGGCCGGCGAAGGTTGGCGAGGGGATCACCC  
TGACTAGGGAGGGTATATTGTCGTCTCCGCCGGCCGCTTCCAACCGCTCCCCTAGTGGG  
2850 EAG1 XMA3,

2882 ProSerValAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuLysAlaThrCys  
CCCTCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCTCAAGGCAACTTGC  
GGGAGACACCGGTCGAGGAGCCGATCGGTCGATAGGCGAGGTAGAGAGTTCGGTTGAACG  
2889 BALI, 2903 NHEI,

FIG. 17E



2942 ThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrpArgGln  
ACCGCTAACCATGACTCCCCCTGATGCTGAGCTCATAGAGGCCAACCTCCTATGGAGGCAG  
TGGCGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGGTTGGAGGATACCTCCGTC  
^ ^

2966 ESP1, 2969 SACI,

3002 GluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeuAspSer  
GAGATGGGCGGCAACATCACCAGGGTTGAGTCAGAAAACAAAGTGGTGATTCTGGACTCC  
CTCTACCCGCCGTTGTAGTGGTCCCAACTCAGTCTTTTGTTCACCACTAAGACCTGAGG

3062 PheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGluIleLeu  
TTCGATCCGCTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGCAGAAATCCTG  
AAGCTAGGCGAACACCGCCTCCTCCTGCTCGCCCTCTAGAGGCATGGGCGTCTTTAGGAC  
^

3096 BGL2,

3122 ArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyrAsnPro  
CGGAAGTCTCGGAGATTGCCCCAGGCCCTGCCGTTTGGGCGCGGCCGACTATAACCCC  
GCCTTCAGAGCCTCTAAGCGGGTCCGGGACGGGCAAACCCGCGCCGGCCTGATATTGGGG  
^ ^

3143 ALWN1, 3164 EAG1 XMA3,

3182 ProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGlyCysPro  
CCGCTAGTGGAGACGTGGAAGCCCGACTACGAACCACCTGTGGTCCATGGCTGCCCCG  
GGCGATCACCTCTGCACCTTTTTCGGGCTGATGCTTGGTGGACACCAAGTACCGACGGGC  
^ ^

3217 HGIE2, 3229 NCOI,

3242 LeuProProProLysSerProProValProProProArgLysLysArgThrValValLeu  
CTTCCACCTCCAAAGTCCCCCTCCTGTGCCTCCGCCTCGGAAGAAGCGGACGGTGGTCCTC  
GAAGGTGGAGGTTTCAGGGGAGGACACGGAGGCGGAGCCTTCTTCGCCTGCCACCAGGAG

3302 ThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGlySerSer  
ACTGAATCAACCCTATCTACTGCCCTGGCCGAGCTCGCCACCAGAGCTTTGGCAGCTCC  
TGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGGTCTTCGAAACCGTCGAGG  
^ ^

3332 SACI, 3346 HIND3,

3362 SerThrSerGlyIleThrGlyAspAsnThrThrThrSerSerGluProAlaProSerGly  
TCAACTTCCGGCATTACGGGCGACAATACGACAACATCCTCTGAGCCCGCCCTTCTGGC  
AGTTGAAGGCCGTAATGCCCGCTGTTATGCTGTTGTAGGAGACTCGGGCGGGGAAGACCG

3422 CysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGlyGluPro  
TGCCCCCGGACTCCGACGCTGAGTCTTATCCTCCATGCCCCCCTGGAGGGGGAGCCT  
ACGGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGGACCTCCCCCTCGGA  
^

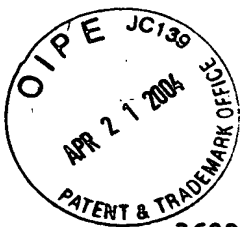
3437 EAM11051,

3482 GlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsnAlaGlu  
GGGGATCCGGATCTTAGCGACGGGTCATGGTCAACGGTCAGTAGTGAGGCCAACGCGGAG  
CCCTAGGCCTAGAATCGCTGCCAGTACCAGTTGCCAGTCATCACTCCGTTGCGCCTC  
^ ^ ^

3484 BAMHI, 3485 BSAB1, 3487 BSPE1,

3542 AspValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuValThrProCysAla  
GATGTCGTGTGCTGCTCAATGTCTTACTCTTGGACAGGCGCACTCGTCACCCCGTGCGCC  
CTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCGCGTGAGCAGTGGGGCACGCGG

FIG. 17F



3602 AlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsn  
GCGGAAGAACAGAACTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTCACCACAAT  
CGCCTTCTTGCTTTGACGGGTAGTTACGTGATTGCTTGAGCAACGATGCAGTGGTGTTA  
3611 ALWN1, 3655 PFLM1,  
3662 LeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAsp  
TTGGTGTATTCCACCACCTCACGCAGTGCTTGCCAAAGGCAGAAGAAAGTCACATTTGAC  
AACCACATAAGGTGGTGGAGTCCGTCACGAACGGTTTCCGTCTTCTTTCAGTGTAAGCTG  
3681 DRA3,  
3722 ArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAlaAlaAla  
AGACTGCAAGTTCTGGACAGCCATTACCAGGACGTAAGGAGGTTAAAGCAGCGGCG  
TCTGACGTTCAAGACCTGTCGGTAATGGTCTGCATGAGTTCCTCCAATTTTCGTCGCCGC  
3782 SerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrProProHis  
TCAAAAGTGAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGACGCTGACGCCCCACAC  
AGTTTTCACTTCCGATTGAACGATAGGCATCTCCTTGAACGTCGGACTGCGGGGTGTG  
3816 HIND3,  
3842 SerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAla  
TCAGCCAAATCCAAGTTTGGTTATGGGGCAAAGACGTCCTGTCATGCCAGAAAGGCC  
AGTCGGTTTAGGTTCAAACCAATACCCCGTTTTCTGCAGGCAACGGTACGGTCTTTCCGG  
3875 AAT2, 3890 BGLI,  
3902 ValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrProIleAsp  
GTAACCCACATCAACTCCGTGTGGAAAGACCTTCTGGAAGACAATGTAACACCAATAGAC  
CATTGGGTGTAGTTGAGGCACACCTTTCTGGAAGACCTTCTGTTACATTGTGGTTATCTG  
3962 ThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGlyArgLys  
ACTACCATCATGGCTAAGAACGAGGTTTTCTGCGTTACGCTGAGAAGGGGGGTCTAAG  
TGATGGTAGTACCGATTCTTGCTCCAAAAGACGCAAGTCGGACTCTTCCCCCAGCATT  
4022 ProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMetAlaLeu  
CCAGCTCGTCTCATCGTGTTCCTCGATCTGGGCGTGCGCGTGTGCGAAAAGATGGCTTTG  
GGTCGAGCAGAGTAGCACAAAGGGGCTAGACCCGCACGCGCACACGCTTTTCTACCGAAC  
4082 TyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPheGlnTyr  
TACGACGTGGTTACAAAGCTCCCTTGGCCGTGATGGGAAGCTCCTACGGATTCCAATAC  
ATGCTGCACCAATGTTTCGAGGGGAACCGGCACTACCTTCGAGGATGCCTAAGGTTATG  
4142 SerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThrProMet  
TCACCAGGACAGCGGGTTGAATTCCTCGTGCAAGCGTGGAAGTCCAAGAAAACCCCAATG  
AGTGGTCTGTGCCCCAACTTAAGGAGCACGTTGCACCTTCAGGTTCTTTTGGGGTTAC  
4160 ECORI,  
4202 GlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIleArgThr  
GGGTTCTCGTATGATACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATCCGTACG  
CCCAAGAGCATACTATGGGCGACGAAACTAGGTGTGCTGAGTCTCGCTGTAGGCATGC

FIG. 17G



4229 DRD1, 4236 ALWN1,

4262 GluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIleLysSer  
GAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCAAGCCGCGTGGCCATCAAGTCC  
CTCCTCCGTTAGATGGTTACAACACTGGAGCTGGGGGTTCTGGGCGCACCGGTAGTTCAGG  
^ ^

4301 BGLI, 4308 BALI,

4322 LeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsnCysGly  
CTCACCAGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAGAACTGCGGC  
GAGTGGCTCTCCGAAATACAACCCCGGGAGAATGGTTAAGTTCCCCCTCTTGACGCCG  
^

4345 APAI,

4382 TyrArgArgCysArgAlaSerGlyValLeuThrThrSerCysGlyAsnThrLeuThrCys  
TATCGCAGGTGCCGCGCGAGCGGCGTACTGACAACTAGCTGTGGTAACACCTCACTTGC  
ATAGCGTCCACGGCGCGCTCGCCGCATGACTGTTGATCGACACCATTTGTGGGAGTGAACG

4442 TyrIleLysAlaArgAlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMetLeuVal  
TACATCAAGGCCCGGGCAGCCTGTCTGAGCCGCGAGGCTCCAGGACTGCACCATGCTCGTG  
ATGTAGTTCCGGGCCCGTCTCGGACAGCTCGGCGTCCCGAGGTCCTGACGTGGTACGACAC  
^

4452 SMAI XMAI,

4502 CysGlyAspAspLeuValValIleCysGluSerAlaGlyValGlnGluAspAlaAlaSer  
TGTGGCGACGACTTAGTCGTTATCTGTGAAAGCGCGGGGGTCCAGGAGGACGCGCGAGC  
ACACCGCTGCTGAATCAGCAATAGACACTTTCGCGCCCCCAGGTCCTCTGCGCCGCTCG  
^ ^

4508 DRD1, 4511 TTH3I,

4562 LeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProGln  
CTGAGAGCCTTCACGGAGGCTATGACCAGGTACTCCGCCCCCTGGGGACCCCCACAA  
GACTCTCGGAAGTGCCTCCGATACTGGTCCATGAGGCGGGGGGACCCCTGGGGGGTGT  
^

4622 ProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaHisAsp  
CCAGAATACGACTTGGAGCTCATAACATCATGCTCCTCCAACGTGTCTAGTCGCCCACGAC  
GGTCTTATGCTGAACCTCGAGTATTGTAGTACGAGGAGGTTGCACAGTCAGCGGGTGCTG  
^

4637 SACI,

4682 GlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAlaArgAla  
GGCGCTGGAAAGAGGGTCTACTACCTACCCGTGACCCTACAACCCCCCTCGCGAGAGCT  
CCGCGACCTTTCTCCAGATGATGGAGTGGGCACTGGGATGTTGGGGGGAGCGCTCTCGA  
^

4731 NRUI,

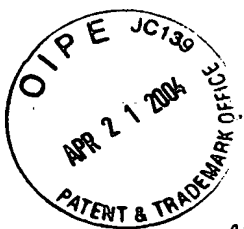
4742 AlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIleMetPhe  
GCGTGGGAGACAGCAAGACACACTCCAGTCAATTCCTGGCTAGGCAACATAATCATGTTT  
CGCACCTCTGTCTGTTCTGTGTGAGGTGAGTTAAGGACCGATCCGTTGTATTAGTACAAA  
^

4802 AlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePheSerValLeuIleAla  
GCCCCACACTGTGGGCGAGGATGATACTGATGACCCATTTCTTTAGCGTCTTATAGCC  
CGGGGGTGTGACACCCGCTCCTACTATGACTACTGGGTAAAGAAATCGCAGGAATATCGG  
^ ^

4806 PFLM1, 4807 DRA3,

ArgAspGlnLeuGluGlnAlaLeuAspCysGluIle TyrGlyAlaCysTyrSerIleGlu

FIG. 17H



4862 AGGGACCAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGGCCTGCTACTCCATAGAA  
TCCCTGGTCGAACTTGTCCGGGAGCTAACGCTCTAGATGCCCCGGACGATGAGGTATCTT  
4893 BGL2,  
ProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSerLeuHis  
4922 CCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTTCACTCCAC  
GGTGACCTAGATGGAGGTTAGTAAGTTTCTGAGGTACCGGAGTCGCGTAAAAGTGAGGTG  
4954 NCOI,  
SerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGlyValPro  
4982 AGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGCATGCCTCAGAAAACTTGGGGTACCG  
TCAATGAGAGGTCCACTTTAGTTATCCACCGCGGTACGGAGTCTTTTGAACCCCATGGC  
5015 SPHI, 5035 KPNI,  
ProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAlaArgGly  
5042 CCCTTGGGAGCTTGGAGACACCGGGCCCGGAGCGTCCGCGCTAGGCTTCTGGCCAGAGGA  
GGGAACGCTCGAACCTCTGTGGCCCGGGCCTCGCAGGCGCGATCCGAAGACCGGTCTCCT  
5064 APAI, 5091 BALI,  
GlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLys  
5102 GGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAAGCTCAA  
CCGTCCCGACGGTATACACCGTTTCATGGAGAAGTTGACCCGTCATTCTTGTTCGAGTTT  
5113 NDEI,  
LeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAlaGlyTyr  
5162 CTCACTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCCGGCTGGTTTACGGCTGGCTAC  
GAGTGAGGTTATCGCCGGCGACCGGTGCGACCTGAACAGGCCGACCAAGTGCCGACCGATG  
5174 NOTI, 5175 EAG1 XMA3, 5182 BALI, 5186 PVU2,  
SerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCys  
5222 AGCGGGGGAGACATTTATCACAGCGTGTCTCATGCCCGGCCCGCTGGATCTGGTTTTGC  
TCGCCCCCTCTGTAAATAGTGTGCGACAGAGTACGGGGCCGGGGCGACCTAGACCAAACG  
5240 DRA3,  
LeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgMetSerThrAsn  
5282 CTACTCCTGCTTGCTGCAGGGGTAGGCATCTACCTCCTCCCAACCGAATGAGCACGAAT  
GATGAGGACGAACGACGTCCCCATCCGTAGATGGAGGAGGGGTGGCTTACTCGTGCTTA  
5295 PSTI,  
ProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGlnAspValLysPhe  
5342 CCTAAACCTCAAAGAAAGACCAAAACGTAACACCAACCGGGCGGCCGAGGACGTCAAGTTC  
GGATTTGGAGTTTCTTTCTGGTTTGCATTGTGGTTGGCCGCGCGGCGTCTGCAGTTCAAG  
5380 NOTI, 5381 EAG1 XMA3, 5390 AAT2, 5401 SMAI XMAI,  
ProGlyGlyGlyGlnIleValGlyGlyValTyrLeuLeuProArgArgGlyProArgLeu  
5402 CCGGGTGGCGGTGAGATCGTTGGTGGAGTTTACTTGTGTCGCGCGAGGGGCCCTAGATTG  
GGCCACCGCCAGTCTAGCAACCACCTCAAATGAACAACGGCGCGTCCCGGGATCTAAC

FIG. 17I



5449 APAI,

5462 GlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGlyArgArgGlnPro  
GGTGTGCGCGCGACGAGAAAGACTTCCGAGCGGTGCAACCTCGAGGTAGACGTCAGCCT  
CCACACGCGCGCTGCTCTTTCTGAAGGCTCGCCAGCGTTGGAGCTCCATCTGCAGTCGGA  
5467 BSSH2, 5478 XMNI, 5502 XHOI, 5511 AAT2,

5522 IleProLysAlaArgArgProGluGlyArgThrTrpAlaGlnProGlyTyrProTrpPro  
ATCCCCAAGGCTCGTCGGCCCCGAGGGCAGGACCTGGGCTCAGCCCGGGTACCCTTGGCCC  
TAGGGGTTCCGAGCAGCCGGGCTCCCGTCTGGACCCGAGTCGGGCCCCATGGGAACCGGG  
5548 ALWN1, 5558 ESP1, 5564 SMAI XMAI, 5568 KPNI,

5582 LeuTyrGlyAsnGluGlyCysGlyTrpAlaGlyTrpLeuLeuSerProArgGlySerArg  
CTCTATGGCAATGAGGGCTGCGGGTGGGCGGGATGGCTCCTGTCTCCCCGTGGCTCTCGG  
GAGATACCGTTACTCCCCGACGCCACCCGCCCTACCGAGGACAGAGGGGCACCGAGAGCC

5642 ProSerTrpGlyProThrAspProArgArgArgSerArgAsnLeuGlyLysOC AM  
CCTAGCTGGGGCCCCACAGACCCCGGCGTAGGTCGCGCAATTTGGGTAAGTAATAGTCG  
GGATCGACCCCGGGGTGTCTGGGGGCCGCATCCAGCGCGTTAAACCCATTTCATTATCAGC  
5650 APAI, 5698 SALI,

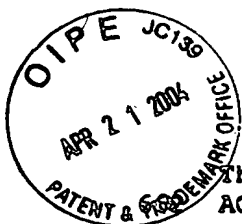
5702 AC  
TG

FIG. 17J



MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsn  
2 AGCTTACAAAACAAAATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTACTCAAC  
TCGAATGTTTTGTTTTACCAGCGTATACGTCGAGTCCCGATATTCCACGATCATGAGTTG  
^ ^ ^  
1 HIND3, 24 NDEI, 52 SCAI,  
  
ProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGlyIleAsp  
62 CCCTCTGTTGCTGCAACACTGGGCTTTGGTGCTTACATGTCCAAGGCTCATGGGATCGAT  
GGGAGACAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCCTAGCTA  
^  
116 CLAI,  
  
ProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyrSerThr  
122 CCTAACATCAGGACCGGGTGAGAACAATTACCACTGGCAGCCCCATCAGTACTCCACC  
GGATTGTAGTCCTGGCCCCACTCTTGTTAATGGTGACCGTCGGGGTAGTGCATGAGGTGG  
  
TyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIleIleCys  
182 TACGGCAAGTTCCTTGCCGACGGCGGGTGCTCGGGGGGCGCTTATGACATAATAATTTGT  
ATGCCGTTCAAGGAACGGGTGCCGCCACGAGCCCCCGCGAATACTGTATTATTAAACA  
  
AspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGln  
242 GACGAGTGCCACTCCACGGATGCCACATCCATCTTGGGCATTGGCACTGTCCTTGACCAA  
CTGCTCACGGTGAGGTGCCTACGGTGTAGGTAGAACCCGTAACCGTGACAGGAAGTGGT  
  
AlaGluThrAlaGlyAlaArgLeuValValLeuAlaThrAlaThrProProGlySerVal  
302 GCAGAGACTGCGGGGGCGAGACTGGTTGTGCTCGCCACCGCCACCCCTCCGGGCTCCGTC  
CGTCTCTGACGCCCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCGAGGCAG  
^  
303 ALWN1,  
  
ThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIleProPhe  
362 ACTGTGCCCCATCCCAACATCGAGGAGGTTGCTCTGTCCACCACCGGAGAGATCCCTTTT  
TGACACGGGGTAGGGTTGTAGCTCCTCCAACGAGACAGGTGGTGGCCTCTCTAGGGAAAA  
  
TyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePheCysHis  
422 TACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGAGACATCTCATCTTCTGTGTCAT  
ATGCCGTTCCGATAGGGGGAGCTTCATTAGTTCCCCCCTCTGTAGAGTAGAAGACAGTA  
  
SerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsnAlaVal  
482 TCAAAGAAGAAGTGCAGCAACTCGCCGCAAAGCTGGTCGCATTGGGCATCAATGCCGTG  
AGTTTCTTCTTCACGCTGCTTGAGCGGCGTTTCGACCAGCGTAACCCGTAGTTACGGCAC  
  
AlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValValValVal  
542 GCCTACTACCGCGGTCTTGACGTGTCCGTCATCCCGACCAGCGCGCATGTTGTGTCGTCGTG  
CGGATGATGGCGCCAGAACTGCACAGGCAGTAGGGCTGGTCGCCGCTACAACAGCAGCAC  
^ ^ ^  
550 SAC2, 560 DRD1,  
  
AlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsn  
602 GCAACCGATGCCCTCATGACCGGCTATACCGGCGACTTCGACTCGGTGATAGACTGCAAT  
CGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTGACGTTA  
^  
615 BSPH1,

FIG. 18A



ThrCysValThrGlnThrValAspPheSerLeuAspPr ThrPheThrIleGluThrIle  
ACGTGTGTCAACCAGACAGTCGATTTACGCCCTTGACCCTACCTTCACCATTGAGACAATC  
TGCACACAGTGGGTCTGTGCTAGCTAAAGTCGGAACCTGGGATGGAAGTGGTAACTCTGTTAG

722 ThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLys  
ACGCTCCCCCAAGATGCTGTCTCCCGCACTCAACGTCGGGGCAGGACTGGCAGGGGGAAG  
TGCGAGGGGGTTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCTGACCGTCCCCCTTC

782 ProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSer  
CCAGGCATCTACAGATTTGTGGCACCAGGGGAGCGCCCCCTCCGGCATGTTGACTCGTCC  
GGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAGCTGAGCAGG  
^ ^

816 BGLI, 833 DRD1,

842 ValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThr  
GTCCTCTGTGAGTGCTATGACGCAGGCTGTGCTTGGTATGAGCTCACGCCCGCCGAGACT  
CAGGAGACACTCACGATACTGCGTCCGACACGAACCATACTCGAGTGCGGGCGGCTCTGA  
^

881 SACI,

902 ThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeu  
ACAGTTAGGCTACGAGCGTACATGAACACCCCGGGGCTTCCCGTGTGCCAGGACCATCTT  
TGTCATCCGATGCTCGCATGTACTTGTGGGCCCCGAAGGGCACACGGTCTTGGTAGAA  
^

931 SMAI XMAI,

962 GluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGln  
GAATTTTGGGAGGGCGTCTTTACAGGCCTCACTCATATAGATGCCCACTTTCTATCCAG  
CTTAAAACCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGATAGGGTC  
^

985 STUI,

1022 ThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAla  
ACAAAGCAGAGTGGGGAGAACCTTCCTTACCTGGTAGCGTACCAAGCCACCGTGTGCGCT  
TGTTTCGTCTACCCCTCTTGGAAGGAATGGACCATCGCATGGTTCCGGTGGCACACGCGA  
^

1069 DRA3,

1082 ArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysLeuIleArgLeuLys  
AGGGCTCAAGCCCCCTCCCCATCGTGGGACCAGATGTGGAAGTGTGTTGATTGCGCTCAAG  
TCCCCAGTTCCGGGAGGGGGTAGCACCTGGTCTACACCTTCACAACTAAGCGGAGTTC

1142 ProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsnGluIle  
CCCACCCTCCATGGGCCAACACCCCTGCTATACAGACTGGGCGCTGTTTCAAGATGAAATC  
GGGTGGGAGGTACCCGGTTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTACTTTAG  
^

1150 NCOI,

1202 ThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeuGluVal  
ACCCTGACGCACCCAGTCACCAAATACATCATGACATGCATGTCGGCCGACCTGGAGGTC  
TGGGACTGCGTGGGTGAGTGGTTTATGTAGTACTGTACGTACAGCCGGCTGGACCTCCAG  
^ ^ ^ ^ ^

1230 BSPH1, 1234 DRD1, 1237 AVA3, 1245 EAG1 XMA3, 1250 DRD1,

1262 ValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyrCysLeu  
GTCACGAGCACCTGGGTGCTCGTTGGCGGCGTCCTGGCTGCTTTGGCCGCGTATTGCCTG

FIG. 18B



CAGTGCTCGTGGACCCACGAGCAACCGCCGCAGGACCGACGAAACCGGCGCATAACGGAC

1322 SerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAlaIleIle  
TCAACAGGCTGCGTGGTCATAGTGGGCAGGGTCGTCTTGTCCGGAAGCCGGCAATCATA  
AGTTGTCCGACGCACCAGTATCACCCGTCCCAGCAGAACAGGCCCTTCGGCCGTTAGTAT  
^

1369 NAEI,

1382 ProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGlnHisLeu  
CCTGACAGGGAAGTCCTCTACCGAGAGTTCGATGAGATGGAAGAGTGCTCTCAGCACTTA  
GGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCACGAGAGTCGTGAAT  
^

1385 DRD1,

1442 ProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeuGlyLeu  
CCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGCAGAAGGCCCTCGGCCTC  
GGCATGTAGCTCGTTCCTACTACGAGCGGCTCGTCAAGTTCGTCTTCGGGAGCCGGAG

1502 LeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsnTrpGln  
CTGCAGACCGCGTCCCGTCAGGCAGAGTTATCGCCCCTGCTGTCCAGACCAACTGGCAA  
GACGTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTGACCGTT  
^ ^

1502 PSTI, 1507 TTH3I,

1562 LysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeu  
AAACTCGAGACCTTCTGGGCGAAGCATATGTGGAACCTTCATCAGTGGGATACAATACTTG  
TTTGAGCTCTGGAAGACCCGCTTCGTATACACCTTGAAGTAGTCACCCTATGTTATGAAC  
^ ^

1565 XHOI, 1586 NDEI,

1622 AlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPheThrAla  
GCGGGCTTGTCACGCTGCCTGGTAACCCCGCCATTGCTTCATTGATGGCTTTTACAGCT  
CGCCCGAACAGTTGCGACGGACCATTTGGGGCGGTAACGAAGTAACCTACCGAAAATGTCTGA  
^ ^

1643 BSTE2, 1677 ALWN1 PVU2,

1682 AlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGlyGlyTrp  
GCTGTCACCAGCCCACTAACCCTAGCCAAACCCTCCTCTTCAACATATTGGGGGGGTGG  
CGACAGTGGTCGGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCCCCACC

1742 ValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGlyLeuAla  
GTGGCTGCCCAGCTCGCCGCCCCCGGTGCCGCTACTGCCTTTGTGGGCGCTGGCTTAGCT  
CACCGACGGGTCGAGCGGGCGGGGGCCACGGCGATGACGGAAACACCCGCGACCGAATCGA  
^

1794 ESI1,

1802 GlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyr  
GGCGCCGCCATCGGCAGTGTTGGACTGGGGAAGGTCCTCATAGACATCCTTGCAGGGTAT  
CCGCGGCGGTAGCCGTCACAACCTGACCCCTTCCAGGAGTATCTGTAGGAACGTCCCATA  
^

1802 KAS1 NARI,

1862 GlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluValProSer  
GGCGCGGGCGTGGCGGGAGCTCTTGTGGCATTCAAGATCATGAGCGGTGAGGTCCCCTCC  
CCGCGCCCGCACCGCCCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAGGGGAGG  
^ ^

1878 SACI, 1899 BSPH1,

FIG. 18C



1922 ThrGluAspLeuValAsnLeuLeuPr AlaIleLeuSerPr GlyAlaLeuValValGly  
ACGGAGGACCTGGTCAATCTACTGCCCGCCATCCTCTCGCCCGGAGCCCTCGTAGTCGGC  
TGCCCTCCTGGACCAGTTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCATCAGCCG  
1928 TTH3I,  
1982 ValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrp  
GTGGTCTGTGCAGCAATACTGCCCGGCACGTTGGCCCGGGCGAGGGGGCAGTGCAGTGG  
CACCAGACACGTCGTTATGACGCGGCCGTGCAACCGGGCCGCTCCCCCGTCACGTCACC  
2004 NAEI, 2017 SMAI XMAI,  
2042 MetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrVal  
ATGAACCGGCTGATAGCCTTCGCCTCCCGGGGAACCATGTTTCCCCACGCACTACGTG  
TACTTGCCGACTATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGGTGCCTGATGCAC  
2067 SMAI XMAI, 2093 DRA3,  
2102 ProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGln  
CCGGAGAGCGATGCAGCTGCCCGCGTCACTGCCATACTCAGCAGCCTCACTGTAACCCAG  
GGCCTCTCGCTACGTCGACGGGCGCAGTGACGGTATGAGTCGTCGGAGTGACATTGGGTC  
2115 PVU2, 2159 ALWN1,  
2162 LeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSerGlySer  
CTCCTGAGGCGACTGCACCACTGGATAAGCTCGGAGTGTACCACTCCATGCTCCGGTTCC  
GAGGACTCCGCTGACGTGGTCACCTATTCGAGCCTCACATGGTGAGGTACGAGGCCAAGG  
2164 MST2, 2220 ECON1,  
2222 TrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeu  
TGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCGACTTTAAGACCTGGCTA  
ACCGATTCCCTGTAGACCCTGACCTATACGCTCCACAACCTCGCTGAAATTCTGGACCGAT  
2282 LysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyr  
AAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCTTTGTGTCTGCCAGCGCGGGTAT  
TTTCGATTGAGTACGGTGTGACGGACCCTAGGGGAAACACAGGACGGTTCGCGCCCAT  
2285 ESP1, 2300 PVU2, 2310 BAMHI,  
2342 LysGlyValTrpArgGlyAspGlyIleMetHisThrArgCysHisCysGlyAlaGluIle  
AAGGGGGTCTGGCGAGGGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCTGAGATC  
TTCCCCAGACCGCTCCCCTGCCGTAGTACGTGTGAGCGACGGTGACACCTCGACTCTAG  
2402 ThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArgAsnMet  
ACTGGACATGTCAAAAACGGGACGATGAGGATCGTCGGTCCTAGGACCTGCAGGAACATG  
TGACCTGTACAGTTTTTGCCTGCTACTCCTAGCAGCCAGGATCCTGGACGTCCTTGATC  
2425 BSAB1, 2441 AVR2, 2448 SSE83871, 2449 PSTI,  
2462 TrpSerGlyThrPheProIleAsnAlaTyrThrThrGlyProCysThrProLeuProAla  
TGGAGTGGGACCTTCCCATTAAATGCCTACACCACGGGCCCCCTGTACCCCCCTTCTGCG  
ACCTCACCTGGAAGGGGTAATTACGGATGTGGTGCCCGGGGACATGGGGGGAAGGACGC  
2480 ASE1, 2497 APAI,

FIG. 18D



2522 ProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIleArgGln  
CCGAAC TACACG TTCGCGCTATGGAGGGTGTCTGCAGAGGAATACGTGGAGATAAGGCAG  
GGCTTGATGTGCAAGCGCGATACCTCCCACAGACGTCTCCTTATGCACCTCTATTCCGTC  
^

2553 PSTI,

2582 ValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysProCysGln  
GTGGGGGACTTCCACTACGTGACGGGTATGACTACTGACAATCTTAAATGCCCGTGCCAG  
CACCCTGAAGGTGATGCACTGCCCATCTGATGACTGTTAGAATTACGGGCACGGTC  
^

2594 DRA3,

2642 ValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPheAlaPro  
GTCCCATCGCCGAATTTTTCACAGAATTGGACGGGGTGC GCCTACATAGGTTTGCGCCC  
CAGGGTAGCGGGCTTAAAAAGTGTCTTAACCTGCCCCACGGGATGTATCCAAACGCGGG

2702 ProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGluTyrPro  
CCCTGCAAGCCCTTGCTGCGGGAGGAGGTATCATTAGAGTAGGACTCCACGAATACCCG  
GGGACGTTCCGGGAACGACGCCCTCCTCCATAGTAAGTCTCATCCTGAGGTGCTTATGGGC  
^

2757 HGIE2,

2762 ValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSerMetLeu  
GTAGGGTCGCAATTACCTTGCGAGCCCGAACCGGACGTGGCCGTGTTGACGTCCATGCTC  
CATCCCAGCGTTAATGGAACGCTCGGGCTTGGCCTGCACCGGCACAAC TGCAGGTACGAG  
^

2809 AAT2,

2822 ThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGlySerPro  
ACTGATCCCTCCCATATAACAGCAGAGGGCGGGCGGAAGGTTGGCGAGGGGGATCACCC  
TGACTAGGGAGGGTATATTGTCGTCTCCGCGGCCGCTTCCAACCGCTCCCTAGTGGG  
^

2850 EAG1 XMA3,

2882 ProSerValAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuLysAlaThrCys  
CCCTCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCTCAAGGCAACTTGC  
GGGAGACACCGGTCGAGGAGCCGATCGGTCGATAGGCGAGGTAGAGAGTTCCGTTGAACG  
^

2889 BALI, 2903 NHEI,

2942 ThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrpArgGln  
ACCGCTAACCATGACTCCCTGATGCTGAGCTCATAGAGGCCAACCTCCTATGGAGGCAG  
TGCGGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGGTTGGAGGATACCTCCGTC  
^ ^

2966 ESP1, 2969 SACI,

3002 GluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeuAspSer  
GAGATGGGCGGCAACATCACCAGGGTTGAGTCAGAAAACAAAGTGGTGATTCTGGACTCC  
CTCTACCCGCCGTTGTAGTGGTCCCAACTCAGTCTTTTGTTCACCACTAAGACCTGAGG

3062 PheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGluIleLeu  
TTCGATCCGCTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGCAGAAATCCTG  
AAGCTAGGCGAACACCGCCTCCTCCTGCTCGCCCTCTAGAGGCATGGGCGTCTTTAGGAC  
^

3096 BGL2,

ArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyrAsnPro

FIG. 18E



3122 CGGAAGTCTCGGAGATTGCCCCAGGCCCTGCCCGTTTGGGCGCGGCCGGACTATAACCCC  
GCCTTCAGAGCCTCTAAGCGGGTCCGGGACGGGCAAACCCGCGCCGGCCTGATATTGGGG  
^ ^

3143 ALWN1, 3164 EAG1 XMA3,

3182 ProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGlyCysPro  
CCGCTAGTGGAGACGTGGAAAAAGCCCCGACTACGAACCACCTGTGGTCCATGGCTGCCCG  
GGCGATCACCTCTGCACCTTTTTTCGGGCTGATGCTTGGTGGACACCAGGTACCGACGGGC  
^ ^

3217 HGIE2, 3229 NCOI,

3242 LeuProProProLysSerProProValProProProArgLysLysArgThrValValLeu  
CTTCCACCTCCAAAGTCCCCTCCTGTGCCTCCGCCTCGGAAGAAGCGGACGGTGGTCCCTC  
GAAGGTGGAGGTTTCAGGGGAGGACACGGAGGCGGAGCCTTCTTCGCCTGCCACCAGGAG

3302 ThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGlySerSer  
ACTGAATCAACCCTATCTACTGCCTTGCCGAGCTCGCCACCAGAAGCTTTGGCAGCTCC  
TGACTTAGTTGGGATAGATGACGGAACCGCTCGAGCGGTGGTCTTCGAAACCGTCGAGG  
^ ^

3332 SACI, 3346 HIND3,

3362 SerThrSerGlyIleThrGlyAspAsnThrThrThrSerSerGluProAlaProSerGly  
TCAACTTCCGGCATTACGGGCGACAATACGACAACATCCTCTGAGCCCCGCCCTTCTGGC  
AGTTGAAGGCCGTAATGCCCGCTGTTATGCTGTTGTAGGAGACTCGGGCGGGGAAGACCG

3422 CysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGlyGluPro  
TGCCCCCCCCGACTCCGACGCTGAGTCCTATTCTCCATGCCCCCCTGGAGGGGGAGCCT  
ACGGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGGACCTCCCCCTCGGA  
^

3437 EAM11051,

3482 GlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsnAlaGlu  
GGGGATCCGGATCTTAGCGACGGGTGATGGTCAACGGTCAGTAGTGAGGCCAACGCGGAG  
CCCCTAGGCCTAGAATCGCTGCCAGTACCAGTTGCCAGTCATCACTCCGTTGCGCCTC  
^^ ^

3484 BAMHI, 3485 BSAB1, 3487 BSPE1,

3542 AspValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuValThrProCysAla  
GATGTGCTGTGCTGCTCAATGTCTTACTCTTGACAGGCGCACTCGTCACCCCGTGCGCC  
CTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCGCGTGAGCAGTGGGGCACGCGG  
^ ^

3589 DRA3, 3600 SAC2,

3602 AlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsn  
GCGGAAGAACAGAACTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTCACCACAAT  
CGCCTTCTTGTCTTTGACGGGTAGTTACGTGATTGCTTGAGCAACGATGCAGTGGTGTTA  
^ ^

3611 ALWN1, 3655 PFLM1,

3662 LeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAsp  
TTGGTGTATTCCACCACCTCAGCAGTGCTTGCCAAAGGCAGAAGAAAGTCACATTTGAC  
AACCACATAAGGTGGTGGAGTGCGTCACGAACGGTTTCGCTTCTTTTCAGTGTAAACTG  
^

3681 DRA3,

3722 ArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAlaAlaAla  
AGACTGCAAGTTCTGGACAGCCATTACCAGGACGTACTCAAGGAGGTTAAAGCAGCGGCG

FIG. 18F



TCTGACGTTCAAGACCTGTCGGTAATGGTCCTGCATGAGTTCCTCCAATTTCTGTCGCCCCG

3782 SerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrProProHis  
TCAAAAGTGAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGCAGCCTGACGCCCCACAC  
AGTTTTCACTTCCGATTGAACGATAGGCATCTCCTTCGAACGTCGGACTGCGGGGTGTG  
^

3816 HIND3,

3842 SerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAla  
TCAGCCAAATCCAAGTTTGGTTATGGGGCAAAGACGTCCGTTGCCATGCCAGAAAGGCC  
AGTCGGTTTAGGTTCAAACCAATACCCCGTTTTCTGCAGGCAACGGTACGGTCTTTCCGG  
^ ^

3875 AAT2, 3890 BGLI,

3902 ValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrProIleAsp  
GTAACCCACATCAACTCCGTGTGGAAAGACCTTCTGGAAGACAATGTAACACCAATAGAC  
CATTGGGTGTAGTTGAGGCACACCTTTCTGGAAGACCTTCTGTTACATTGTGGTTATCTG

3962 ThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGlyArgLys  
ACTACCATCATGGCTAAGAACGAGGTTTTCTGCGTTACAGCCTGAGAAGGGGGGTCTGTAAG  
TGATGGTAGTACCGATTCTTGCTCCAAAAGACGCAAGTCGGACTCTTCCCCCAGCATTC

4022 ProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMetAlaLeu  
CCAGCTCGTCTCATCGTGTTCCTCGATCTGGGCGTGCGCGTGTGCGAAAAGATGGCTTTG  
GGTCGAGCAGAGTAGCACAAAGGGGCTAGACCCGCACGCGCACACGCTTTTCTACCGAAAC

4082 TyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPheGlnTyr  
TACGACGTGGTTACAAAGCTCCCTTGCGCGTGATGGGAAGCTCCTACGGATTCCAATAC  
ATGCTGCACCAATGTTTCGAGGGGAACCGGCACTACCTTCGAGGATGCCTAAGGTTATG

4142 SerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThrProMet  
TCACCAGGACAGCGGTTGAATTCCTCGTGCAAGCGTGGAAGTCCAAGAAAACCCCAATG  
AGTGGTCCTGTCGCCCACTTAAGGAGCACGTTTCGCACCTTCAGGTTCTTTGGGGTTAC  
^

4160 ECORI,

4202 GlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIleArgThr  
GGGTTCTCGTATGATACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATCCGTACG  
CCAAGAGCATACTATGGGCGACGAAACTGAGGTGTCAGTGACTCTCGCTGTAGGCATGC  
^ ^

4229 DRD1, 4236 ALWN1,

4262 GluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIleLysSer  
GAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCAAGCCCGCTGGCCATCAAGTCC  
CTCCTCCGTTAGATGGTTACAACACTGGAGCTGGGGGTTCTGGGCGCACCGGTAGTTTCAGG  
^ ^

4301 BGLI, 4308 BALI,

4322 LeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsnCysGly  
CTCACCAGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAGAAGTGGCGC  
GAGTGGCTCTCCGAAATACAACCCCGGAGAAATGGTTAAGTTCCCCCTCTTGACGCCG  
^

4345 APAI,

4382 TyrArgArgCysArgAlaSerGlyValL uThrThrSerCysGlyAsnThrLeuThrCys  
TATCGCAGGTGCCGCGCGAGCGGCTACTGACAACTAGCTGTGGTAACACCCTCACTTGC  
ATAGCGTCCACGGCGCGCTCGCCGCTGACTGTTGATCGACACCATTTGTGGGAGTGAACG

FIG. 18G



4442 TyrIleLysAlaArgAlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMetLeuVal  
TACATCAAGGCCCGGGCAGCTTGTGCGAGCCGAGGGCTCCAGGACTGCACCATGCTCGTG  
ATGTAGTTCCGGGCGGTCGGACAGCTCGGCGTCCCGAGGTCTGACGTGGTACGAGCAC  
^  
4452 SMAI XMAI,

4502 CysGlyAspAspLeuValValIleCysGluSerAlaGlyValGlnGluAspAlaAlaSer  
TGTGGCGACGACTTAGTCGTTATCTGTGAAAGCGCGGGGGTCCAGGAGGACGCGGCGAGC  
ACACCGCTGCTGAATCAGCAATAGACACTTTCGCGCCCCCAGGTCTCCTGCGCCGCTCG  
^ ^  
4508 DRD1, 4511 TTH3I,

4562 LeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProGln  
CTGAGAGCCTTCACGGAGGCTATGACCAGGTACTCCGCCCCCTGGGGACCCCCACAA  
GACTCTCGGAAGTGCCTCCGATACTGGTCCATGAGGCGGGGGGACCCCTGGGGGGTGT

4622 ProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaHisAsp  
CCAGAATACGACTTGGAGCTCATAACATCATGCTCCTCCAACGTGTGAGTCGCCCCACGAC  
GGTCTTATGCTGAACCTCGAGTATTGTAGTACGAGGAGGTTGCACAGTCAGCGGGTGTG  
^  
4637 SACI,

4682 GlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAlaArgAla  
GGCGCTGGAAAGAGGGTCTACTACCTCACCCGTGACCCTACAACCCCTCGCGAGAGCT  
CCGCGACCTTTCTCCAGATGATGGAGTGGGCACTGGGATGTTGGGGGGAGCGCTCTCGA  
^  
4731 NRUI,

4742 AlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIleMetPhe  
GCGTGGGAGACAGCAAGACACACTCCAGTCAATTCCTGGCTAGGCAACATAATCATGTTT  
CGCACCCCTCTGTCGTTCTGTGTGAGGTGAGTTAAGGACCGATCCGTTGTATTAGTACAAA

4802 AlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePheSerValLeuIleAla  
GCCCCACACTGTGGGCGAGGATGATACTGATGACCCATTTCTTTAGCGTCCTTATAGCC  
CGGGGTGTGACACCCGCTCCTACTATGACTACTGGGTAAAGAAATCGCAGGAATATCGG  
^ ^  
4806 PFLM1, 4807 DRA3,

4862 ArgAspGlnLeuGluGlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSerIleGlu  
AGGGACCAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGGCCTGCTACTCCATAGAA  
TCCCTGGTGAACCTTGTCCGGGAGCTAACGCTCTAGATGCCCCGACGATGAGGTATCTT  
^  
4893 BGL2,

4922 ProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSerLeuHis  
CCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTTCACTCCAC  
GGTGACCTAGATGGAGGTTAGTAAGTTTCTGAGGTACCGGAGTCGCGTAAAAGTGAGGTG  
^  
4954 NCOI,

4982 SerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGlyValPro  
AGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGCATGCCTCAGAAACTTGGGGTACCG  
TCAATGAGAGGTCCACTTTAGTTATCCACCGGCGTACGGAGTCTTTTGAACCCCATGGC  
^ ^  
5015 SPHI, 5035 KPNI,

FIG. 18H



5042 ProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAlaArgGly  
CCCTTGCGAGCTTGGAGACACCGGGCCCGGAGCGTCCGCGCTAGGCTTCTGGCCAGAGGA  
GGGAACGCTCGAACCTCTGTGGCCCGGGCCTCGCAGGCGCGATCCGAAGACCGGTCTCCT  
5064 APAI, 5091 BALI,  
5102 GlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLys  
GGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAAGCTCAAA  
CCGTCCCGACGGTATACACCGTTCATGGAGAAGTTGACCCGTCATTCTTGTTTCGAGTTT  
5113 NDEI,  
5162 LeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAlaGlyTyr  
CTCACTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCCGGCTGGTTCACGGCTGGCTAC  
GAGTGAGGTTATCGCCGGCGACCGGTTCGACCTGAACAGGCCGACCAAGTGCCGACCGATG  
5174 NOTI, 5175 EAGI XMA3, 5182 BALI, 5186 PVU2,  
5222 SerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCys  
AGCGGGGGAGACATTTATCACAGCGTGTCTCATGCCCCGGCCCCGCTGGATCTGGTTTTGC  
TCGCCCCCTCTGTAAATAGTGTGCGACAGAGTACGGGCCGGGGCGACCTAGACCAAAACG  
5240 DRA3,  
5282 LeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgMetSerThrAsn  
CTACTCCTGCTTGCTGCAGGGGTAGGCATCTACCTCCTCCCCAACC GAATGAGCACGAAT  
GATGAGGACGAACGACGTCCCCATCCGTAGATGGAGGAGGGGTGGCTTACTCGTGCTTA  
5295 PSTI,  
5342 ProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGlnAspValLysPhe  
CCTAAACCTCAAAGAAAGACCAAACGTAACACCAACCGGCCGCGCAGGACGTC AAGTTC  
GGATTTGGAGTTTCTTTCTGGTTTGCATTGTGGTTGGCCGCGCGTCTGCAGTTCAAG  
5380 NOTI, 5381 EAGI XMA3, 5390 AAT2, 5401 SMAI XMAI,  
5402 ProGlyGlyGlyGlnIleValGlyGlyValTyrLeuLeuProArgArgGlyProArgLeu  
CCGGGTGGCGGT CAGATCGTTGGTGGAGTTTACTTGTGTCGCGCAGGGGCCCTAGATTG  
GGCCACCGCCAGTCTAGCAACCACCTCAAATGAACAACGGCGCGTCCCCGGGATCTAAC  
5449 APAI,  
5462 GlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGlyArgArgGlnPro  
GGTGTGCGCGCAGACGAGAAAGACTTCCGAGCGGTGCAACCTCGAGGTAGACGTCAGCCT  
CCACACGCGCGCTGCTCTTTCTGAAGGCTCGCCAGCGTTGGAGTCCATCTGCAGTCGGA  
5467 BSSH2, 5478 XMNI, 5502 XHOI, 5511 AAT2,  
5522 IleProLysAlaArgArgProGluGlyArgThrTrpAlaGlnProGlyTyrProTrpPro  
ATCCCCAAGGCTCGTCGGCCCCGAGGGCAGGACCTGGGCTCAGCCGGGTACCCTTGGCCC  
TAGGGGTTCCGAGCAGCCGGGCTCCCGTCTGGACCCGAGTCGGGCCCATGGGAACCGGG  
5548 ALWN1, 5558 ESP1, 5564 SMAI XMAI, 5568 KPNI,  
5582 LeuTyrGlyAsnGluGlyCysGlyTrpAlaGlyTrpLeuLeuSerProArgGlySerArg  
CTCTATGGCAATGAGGGCTGCGGGTGGGCGGGATGGCTCCTGTCTCCCCGTGGCTCTCGG  
GAGATACCGTTACTCCCGACGCCACCCGCCCTACCGAGGACAGAGGGGCACCGAGAGCC

FIG. 18I



5642 ProSerTrpGlyProThrAspProArgArgArgSerArgAsnLeuGlyLysValIleAsp  
CCTAGCTGGGGCCCCACAGACCCCGGCGTAGGTCGCGCAATTTGGGTAAGGTCATCGAT  
GGATCGACCCCGGGGTGTCTGGGGGCCGCATCCAGCGCGTTAAACCCATTCCAGTAGCTA  
5650 APAI, 5696 CLAI,  
5702 ThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProLeuValGlyAlaProLeu  
ACCCTTACGTGCGGCTTCGCCGACCTCATGGGGTACATACCGCTCGTCGGCGCCCTCTT  
TGGGAATGCACGCCGAAGCGGCTGGAGTACCCCATGTATGGCGAGCAGCCGCGGGGAGAA  
5724 HGIE2, 5750 KAS1 NARI, 5756 ECON1,  
5762 GlyGlyAlaAlaArgAlaLeuAlaHisGlyValArgValLeuGluAspGlyValAsnTyr  
GGAGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTTCTGGAAGACGGCGTGAACCTAT  
CCTCCGCGACGGTCCCGGGACCGCGTACCGCAGGCCAAGACCTTCTGCCGCACTTGATA  
5772 BSTXI, 5775 APAI,  
5822 AlaThrGlyAsnLeuProGlyCysSerOC AM  
GCAACAGGGAACCTTCCTGGTTGCTCTTAATAGTCGAC  
CGTTGTCCCTTGGAAGGACCAACGAGAATTATCAGCTG  
5854 SALI,

FIG. 18J

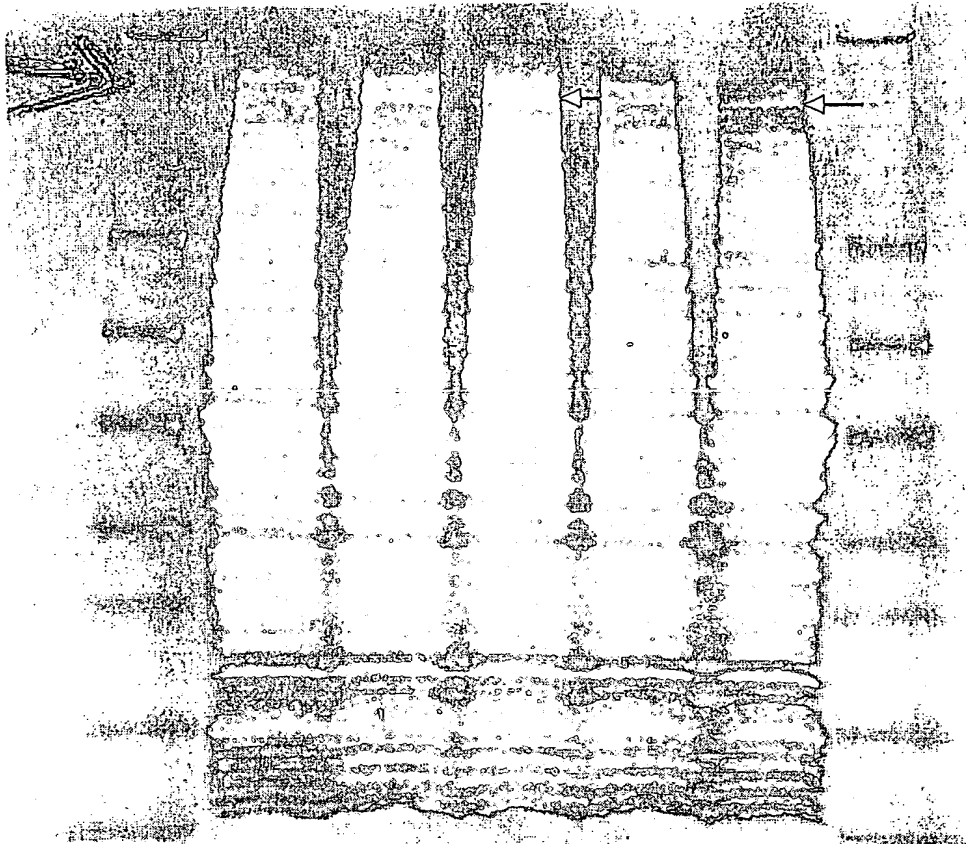


FIG. 19

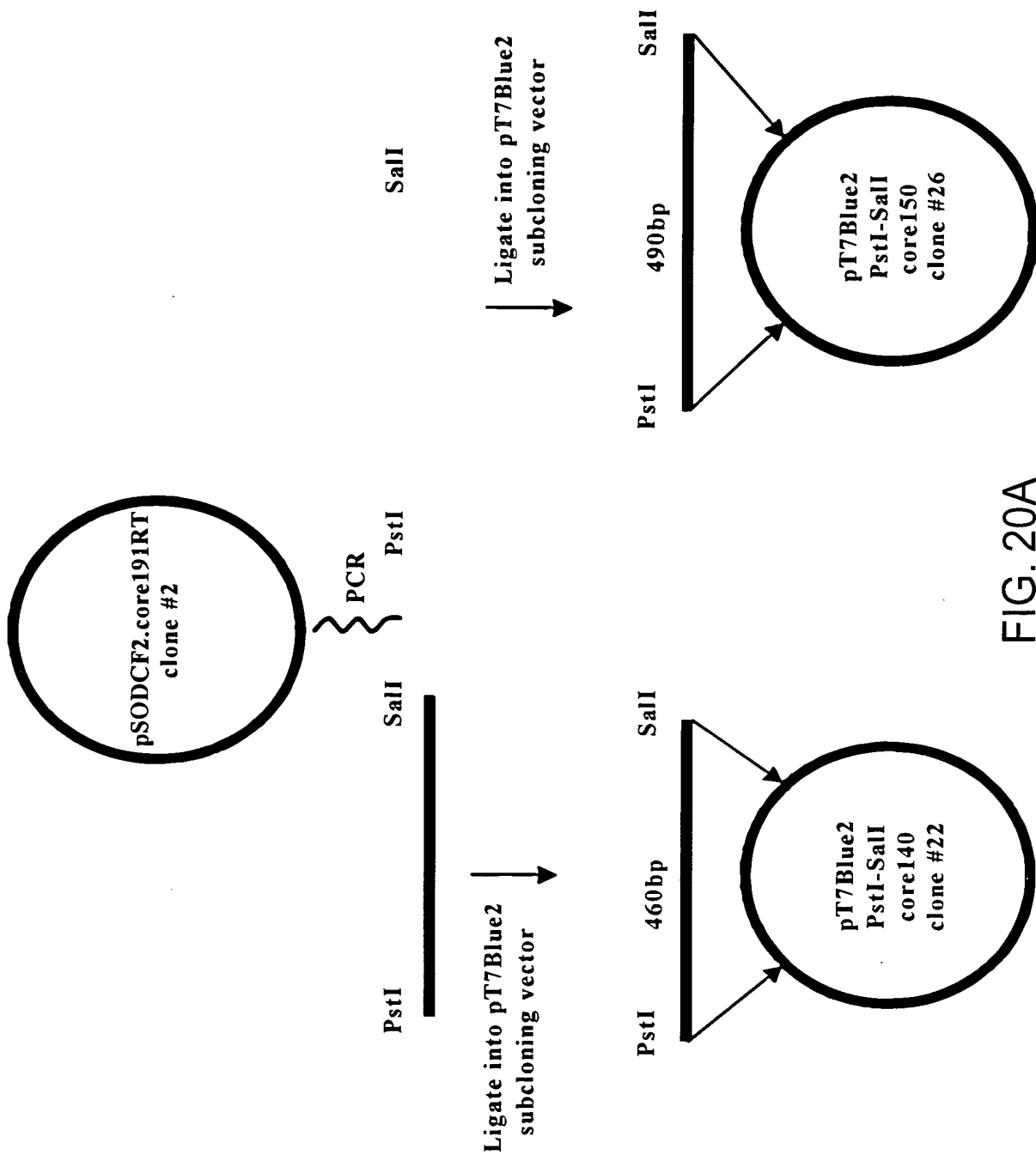


FIG. 20A

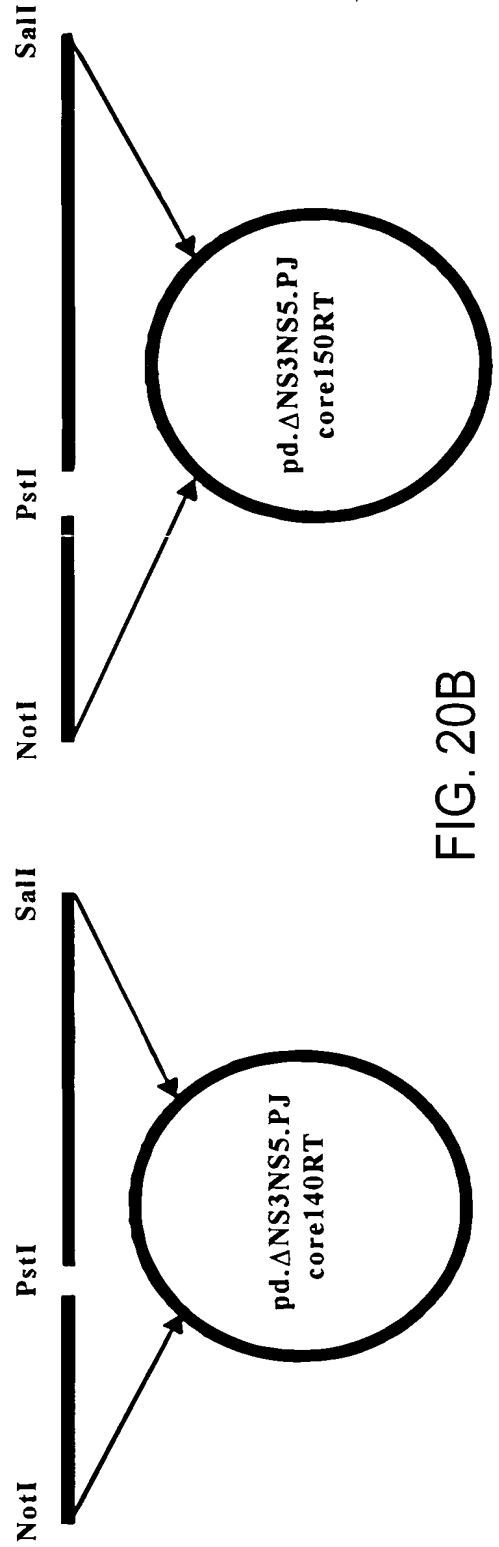
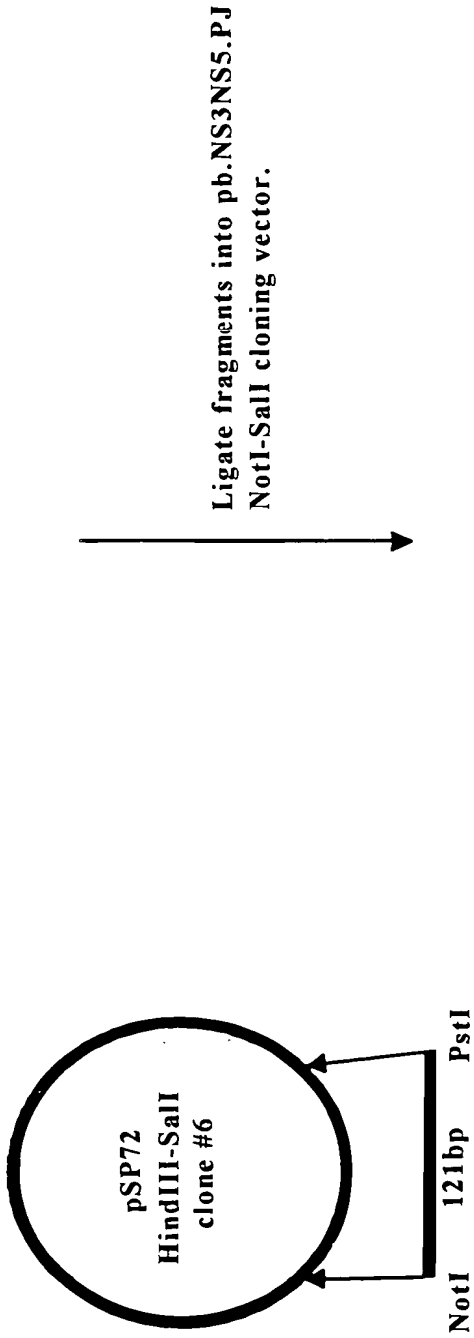


FIG. 20B



MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsn  
2 AGCTTACAAAACAAAATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTACTCAAC  
TCGAATGTTTTGTTTTACCGACGTATACGTCGAGTCCCGATATTCCACGATCATGAGTTG  
^ ^ ^  
1 HIND3, 24 NDEI, 52 SCAI,  
  
ProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGlyIleAsp  
62 CCCTCTGTTGCTGCAACACTGGGCTTTGGTGCTTACATGTCCAAGGCTCATGGGATCGAT  
GGGAGACAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCCCTAGCTA  
^  
116 CLAI,  
  
ProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyrSerThr  
122 CCTAACATCAGGACCGGGGTGAGAACAATTACCACTGGCAGCCCCATCACGTACTCCACC  
GGATTGTAGTCCTGGCCCCACTCTTGTTAATGGTGACCGTCGGGGTAGTGATGAGGTGG  
  
TyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIleIleCys  
182 TACGGCAAGTTCCTTGCCGACGGCGGGTGCTCGGGGGGCGCTTATGACATAATAATTTGT  
ATGCCGTTCAAGGAACGGCTGCCGCCCCACGAGCCCCCGCAATACTGTATTATTAAACA  
  
AspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGln  
242 GACGAGTGCCACTCCACGGATGCCACATCCATCTTGGGCATTGGCACTGTCCTTGACCAA  
CTGCTCACGGTGAGGTGCCTACGGTGTAGGTAGAACCCTAACCGTGACAGGAAGTGGTT  
  
AlaGluThrAlaGlyAlaArgLeuValValLeuAlaThrAlaThrProProGlySerVal  
302 GCAGAGACTGCGGGGGCGAGACTGGTTGTGCTCGCCACCGCCACCCCTCCGGGCTCCGTC  
CGTCTCTGACGCCCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCCGAGGCAG  
^  
303 ALWN1,  
  
ThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIleProPhe  
362 ACTGTGCCCCATCCCAACATCGAGGAGGTTGCTCTGTCCACCACCGGAGAGATCCCTTTT  
TGACACGGGGTAGGGTTGTAGCTCCTCCAACGAGACAGGTGGTGGCCTCTCTAGGGAAAA  
  
TyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePheCysHis  
422 TACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGAGACATCTCATCTTCTGTGTCAT  
ATGCCGTTCCGATAGGGGGAGCTTCATTAGTTCCCCCCTCTGTAGAGTAGAAGACAGTA

FIG. 21A



482 SerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsnAlaVal  
TCAAAGAAGAAGTGCAGCAACTCGCCGCAAAGCTGGTCGCATTGGGCATCAATGCCGTG  
AGTTTCTTCTTCACGCTGCTTGAGCGGCGTTTCGACCAGCGTAACCCGTAGTTACGGCAC

542 AlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValValValVal  
GCCTACTACCGCGGTCTTGACGTGTCCGTATCCCGACCAGCGGCGATGTTGTCTGTCGTG  
CGGATGATGGCGCCAGAACTGCACAGGCAGTAGGGCTGGTCGCCGCTACAACAGCAGCAC  
550 SAC2, 560 DRD1,

602 AlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsn  
GCAACCGATGCCCTCATGACCGGCTATACCGGCGACTTCGACTCGGTGATAGACTGCAAT  
CGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTGACGTTA  
615 BSPH1,

662 ThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGluThrIle  
ACGTGTGTACCCAGACAGTCGATTTTCAGCCTTGACCCTACCTTCACCATTGAGACAATC  
TGCACACAGTGGGTCTGTCTAGCTAAAGTCGGAAGTGGGATGGAAGTGGTAAGTCTGTTAG

722 ThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLys  
ACGCTCCCCAAGATGCTGTCTCCCGCACTCAACGTCGGGGCAGGACTGGCAGGGGGAAG  
TGCGAGGGGGTTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCTGACCGTCCCCCTTC

782 ProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSer  
CCAGGCATCTACAGATTTGTGGCACCGGGGGAGCGCCCTCCGGCATGTTTCGACTCGTCC  
GGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAGCTGAGCAGG  
816 BGLI, 833 DRD1,

842 ValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThr  
GTCCTCTGTGAGTGCTATGACGCAGGCTGTGCTTGGTATGAGCTCACGCCCCGCGAGACT  
CAGGAGACACTCACGATACTGCGTCCGACACGAACCATACTCGAGTGCGGGCGGCTCTGA  
881 SACI,

902 ThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeu  
ACAGTTAGGCTACGAGCGTACATGAACACCCCGGGGCTTCCCGTGTGCCAGGACCATCTT  
TGTCATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCACACGGTCTTGTTAGAA  
931 SMAI XMAI,

962 GluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGln  
GAATTTTGGGAGGGCGTCTTTACAGGCCTCACTCATATAGATGCCCACTTTCTATCCAG  
CTTAAACCCCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGATAGGGTC  
985 STUI,

1022 ThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAla  
ACAAAGCAGAGTGGGGAGAACCTTCTTACCTGGTAGCGTACCAAGCCACCGTGTGCGCT  
TGTTTTCTCTACCCCTCTTGAAGGAATGGACCATCGCATGGTTCGGTGGCACACGCGA  
1069 DRA3,

1082 ArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysLeuIleArgLeuLys  
AGGGCTCAAGCCCCCTCCCCATCGTGGGACCAGATGTGGAAGTGTGTTGATTGCGCTCAAG

FIG. 21B



TCCCGAGTTCGGGGAGGGGGTAGCACCCTGGTCTACACCTTCACAACTAAGCGGAGTTC

1142 ProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsnGluIle  
CCCACCCTCCATGGGCCAACACCCCTGCTATACAGACTGGGCGCTGTTCAAGATGAAATC  
GGGTGGGAGGTACCCGGTTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTACTTTAG  
^

1150 NCOI,

1202 ThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeuGluVal  
ACCCTGACGCACCCAGTCACCAAATACATCATGACATGCATGTCTGGCCGACCTGGAGGTC  
TGGGACTGCGTGGGTGAGTGGTTTATGTAGTACTGTACGTACAGCCGGCTGGACCTCCAG  
^ ^ ^ ^ ^

1230 BSPH1, 1234 DRD1, 1237 AVA3, 1245 EAG1 XMA3, 1250 DRD1,

1262 ValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyrCysLeu  
GTCACGAGCACCTGGGTGCTCGTTGGCGGCGTCTGGCTGCTTTGGCCGCGTATTGCCTG  
CAGTGCTCGTGGACCCACGAGCAACCGCCGACGACCGACGAAACCGGCGCATAACGGAC

1322 SerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAlaIleIle  
TCAACAGGCTGCGTGGTCATAGTGGGCAGGGTCTGCTTGTCCGGGAAGCCGGCAATCATA  
AGTTGTCCGACGCACCAAGTATCACCCGTCCAGCAGAACAGGCCCTTCGGCCGTTAGTAT  
^

1369 NAEI,

1382 ProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGlnHisLeu  
CCTGACAGGGAAGTCCTCTACCGAGAGTTCGATGAGATGGAAGAGTGCTCTCAGCACTTA  
GGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCACGAGAGTCGTGAAT  
^

1385 DRD1,

1442 ProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeuGlyLeu  
CCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGCAGAAGGCCCTCGGCCTC  
GGCATGTAGCTCGTTCCTACTACGAGCGGCTCGTCAAGTTCGTCTTCCGGGAGCCGGAG

1502 LeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsnTrpGln  
CTGCAGACCGCGTCCCGTCAGGCAGAGGTTATCGCCCCTGCTGTCCAGACCAACTGGCAA  
GACGTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTGACCGTT  
^ ^

1502 PSTI, 1507 TTH3I,

1562 LysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeu  
AAACTCGAGACCTTCTGGGCGAAGCATATGTGGAACCTTCATCAGTGGGATACAATACTTG  
TTTGAGCTCTGGAAGACCCGCTTCGTATACACCTTGAAGTAGTCACCCTATGTTATGAAC  
^ ^

1565 XHOI, 1586 NDEI,

1622 AlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPheThrAla  
GCGGGCTTGTCAACGCTGCCTGGTAACCCCGCCATTGCTTCATTGATGGCTTTTACAGCT  
CGCCCGAACAGTTGCGACGGACCATTTGGGGCGGTAACGAAGTAACTACCGAAAATGTCTGA  
^ ^

1643 BSTE2, 1677 ALWN1 PVU2,

1682 AlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGlyGlyTrp  
GCTGTACACGACCCACTAACCACTAGCCAAACCCCTCCTCTTCAACATATTGGGGGGGTGG  
CGACAGTGGTGGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCCCCACC

FIG. 21C



1742 ValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGlyLeuAla  
GTGGCTGCCAGCTCGCCGCCCTCCGGTGCCGCTACTGCCTTTGTGGGCGCTGGCTTAGCT  
CACCGACGGGTCGAGCGGGCGGGGGCCACGGCGATGACGGAAACACCCGCGACCGAATCGA  
1794 ESP1,  
1802 GlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyr  
GGCGCCGCCATCGGCAGTGTGGACTGGGGAAGGTCCTCATAGACATCCTTGCGAGGGTAT  
CCGCGGCGGTAGCCGTCAACCTGACCCCTTCCAGGAGTATCTGTAGGAACGTCCCAT  
1802 KAS1 NARI,  
1862 GlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluValProSer  
GGCGCGGGCGTGGCGGGAGCTCTTGTCGATTCAAGATCATGAGCGGTGAGGTCCCCTCC  
CCGCGCCCGCACCGCCCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAGGGGAGG  
1878 SACI, 1899 BSPH1,  
1922 ThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuValValGly  
ACGGAGGACCTGGTCAATCTACTGCCC GCCATCCTCTCGCCCGGAGCCCTCGTAGTCGGC  
TGCTCTGGACAGTTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCATCAGCCG  
1928 TTH3I,  
1982 ValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrp  
GTGGTCTGTGCAGCAATACTGCGCCGGCACGTTGGCCCGGGCGAGGGGGCAGTGCAAGTGG  
CACCAGACACGTCGTTATGACGCGGCCGTGCAACCGGGCCCGCTCCCCCGTCACGTCACC  
2004 NAEI, 2017 SMAI XMAI,  
2042 MetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrVal  
ATGAACCGGCTGATAGCCTTCGCCTCCCGGGGAACCATGTTTCCCCACGCACTACGTG  
TACTTGCGCGACTATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGGTGCGTGATGCAC  
2067 SMAI XMAI, 2093 DRA3,  
2102 ProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGln  
CCGGAGAGCGATGCAGCTGCCCGCGTCACTGCCATACTCAGCAGCCTCACTGTAACCCAG  
GGCCTCTCGCTACGTCGACGGGCGCAGTGACGGTATGAGTCGTGCGAGTGACATTGGGTC  
2115 PVU2, 2159 ALWN1,  
2162 LeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSerGlySer  
CTCCTGAGGCGACTGCACCAAGTGGATAAGCTCGGAGTGTACCACTCCATGCTCCGGTTCC  
GAGGACTCCGCTGACGTGGTCACCTATTCGAGCCTCACATGGTGAGGTACGAGGCCAAGG  
2164 MST2, 2220 ECON1,  
2222 TrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeu  
TGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCGACTTTAAGACCTGGCTA  
ACCGATTCCCTGTAGACCCTGACCTATACGCTCCACAACCTCGCTGAAATTCTGGACCGAT  
2282 LysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyr  
AAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCTTTGTGTCCTGCCAGCGCGGGTAT  
TTTCGATTGAGTACGGTGTGACGGACCCCTAGGGGAAACACAGGACGGTCGCGCCCAT  
2285 ESP1, 2300 PVU2, 2310 BAMHI,



2342 LysGlyValTrpArgGlyAspGlyIleMetHisThrArgCysHisCysGlyAlaGluIle  
AAGGGGGTCTGGCGAGGGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCTGAGATC  
TTCCCCCAGACCGCTCCCCTGCCGTAGTACGTGTGAGCGACGGTGACACCTCGACTCTAG

2402 ThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArgAsnMet  
ACTGGACATGTCAAAAACGGGACGATGAGGATCGTCGGTCCTAGGACCTGCAGGAACATG  
TGACCTGTACAGTTTTTGGCCCTGCTACTCCTAGCAGCCAGGATCCTGGACGTCTTGTAC  
2425 BSAB1, 2441 AVR2, 2448 SSE83871, 2449 PSTI,

2462 TrpSerGlyThrPheProIleAsnAlaTyrThrThrGlyProCysThrProLeuProAla  
TGGAGTGGGACCTTCCCCATTAATGCCTACACCACGGGCCCCTGTACCCCCCTTCTGCG  
ACCTCACCTGGAAGGGGTAATTACGGATGTGGTGCCCGGGGACATGGGGGAAGGACGC  
2480 ASE1, 2497 APAI,

2522 ProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIleArgGln  
CCGAACTACACGTTCGCGCTATGGAGGGTGTCTGCAGAGGAATACGTGGAGATAAGGCAG  
GGCTTGATGTGCAAGCGCGATACCTCCCACAGACGTCTCCTTATGCACCTCTATTCCGTC  
2553 PSTI,

2582 ValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysProCysGln  
GTGGGGGACTTCCACTACGTGACGGGTATGACTACTGACAATCTTAAATGCCCCGTGCCAG  
CACCCCCCTGAAGGTGATGCACTGCCCATACTGATGACTGTTAGAATTTACGGGCACGGTC  
2594 DRA3,

2642 ValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPheAlaPro  
GTCCCATCGCCGAATTTTTTACAGAATTGGACGGGTGCGCCTACATAGGTTTGCGCCC  
CAGGGTAGCGGGCTTAAAAAGTGCTTTAACCTGCCCCACGGGATGTATCCAAACGCGGG

2702 ProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGluTyrPro  
CCCTGCAAGCCCTTGCTGCGGGAGGAGGTATCATTACAGAGTAGGACTCCACGAATACCCG  
GGGACGTTTCGGGAACGACGCCCTCCTCCATAGTAAGTCTCATCCTGAGGTGCTTATGGGC  
2757 HGIE2,

2762 ValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSerMetLeu  
GTAGGGTTCGCAATTACCTTGCGAGCCCGAACCGGACGTGGCCGTGTTGACGTCCATGCTC  
CATCCAGCGTTAATGGAACGCTCGGGCTTGGCCTGCACCGGCACAACCTGCAGGTACGAG  
2809 AAT2,

2822 ThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGlySerPro  
ACTGATCCCTCCCATATAACAGCAGAGGCGGGCCGGCGAAGGTTGGCGAGGGGATCACCC  
TGACTAGGGAGGGTATATTGTCGTCTCCGCCGGCCGCTTCCAACCGCTCCCCTAGTGGG  
2850 EAG1 XMA3,

2882 ProSerValAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuLysAlaThrCys  
CCCTCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCTCAAGGCAACTTGC  
GGGAGACACCGGTCGAGGAGCCGATCGGTCGATAGGCGAGGTAGAGAGTTCCGTTGAACG  
2889 BALI, 2903 NHEI,

FIG. 21E

Appln. No. 09/721,479  
Replacement Sheet



3589 DRA3, 3600 SAC2,  
AlaGluGluGlnLysLeuPr IleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsn  
3602 GCGGAAGAACAGAAACTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTCAACCACAAT  
CGCCTTCTTGCTTTGACGGGTAGTTACGTGATTCGTTGAGCAACGATGCAGTGGTGTTA  
3611 ALWN1, 3655 PFLM1,  
LeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAsp  
3662 TTGGTGTATTCCACCACCTCACGCAGTGCTTGCCAAAGGCAGAAGAAAGTCACATTTGAC  
AACCACATAAGGTGGTGGAGTGCCTCACGAACGGTTTCCGTCTTCTTTCAGTGTAAGCTG  
3681 DRA3,  
ArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAlaAlaAla  
3722 AGACTGCAAGTTCTGGACAGCCATTACCAGGACGTACTCAAGGAGGTTAAAGCAGCGGCG  
TCTGACGTTCAAGACCTGTCGGTAATGGTCTGCATGAGTTCCTCCAATTTTCGTCGCCCG  
SerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrProProHis  
3782 TCAAAAGTGAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGACGCCTGACGCCCCCACAC  
AGTTTTCACTTCCGATTGAACGATAGGCATCTCCTTCGAACGTGGACTGCGGGGGTG  
3816 HIND3,  
SerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAla  
3842 TCAGCCAAATCCAAGTTTGGTTATGGGGCAAAGACGTCCGTTGCCATGCCAGAAAGGCC  
AGTCGGTTTAGGTTCAAACCAATACCCGTTTTCTGCAGGCAACGGTACGGTCTTTCCGG  
3875 AAT2, 3890 BGLI,  
ValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrProIleAsp  
3902 GTAACCCACATCAACTCCGTGTGGAAAGACCTTCTGGAAGACAATGTAACACCAATAGAC  
CATTGGGTGTAGTTGAGGCACACCTTCTGGAAGACCTTCTGTTACATTGTGGTTATCTG  
ThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGlyArgLys  
3962 ACTACCATCATGGCTAAGAACGAGGTTTTCTGCGTTCAGCCTGAGAAGGGGGGTCGTAAG  
TGATGGTAGTACCGATTCTTGCTCCAAAGACGCAAGTCGGACTCTTCCCCCAGCATTC  
ProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMetAlaLeu  
4022 CCAGCTCGTCTCATCGTGTTCCTCCCGATCTGGGCGTGCGCGTGTGCGAAAAGATGGCTTTG  
GGTCGAGCAGAGTAGCACAAGGGGCTAGACCCGCACGCGCACACGCTTTTCTACCGAAAC  
TyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPheGlnTyr  
4082 TACGACGTGGTTACAAAGCTCCCCTTGGCCGTGATGGGAAGCTCCTACGGATTCCAATAC  
ATGCTGCACCAATGTTTCGAGGGGAACCGGCACTACCCTTCGAGGATGCCTAAGGTTATG  
SerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThrProMet  
4142 TCACCAGGACAGCGGGTTGAATTCTCGTGCAAGCGTGGAAGTCCAAGAAAACCCCAATG  
AGTGGTCTGTGCCCCAAGTTAAGGAGCACGTTTCGCACCTTCAGGTTCTTTTGGGGTTAC  
4160 ECORI,  
GlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIleArgThr  
4202 GGGTTCTCGTATGATACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATCCGTACG  
CCCAAGAGCATACTATGGGCGACGAAACTGAGGTGTCAGTGACTCTCGCTGTAGGCATGC

FIG. 21G



4229 DRD1, 4236 ALWN1,

GluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIleLysSer  
4262 GAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCAAGCCCGCTGGCCATCAAGTCC  
CTCCTCCGTTAGATGGTTACAACACTGGAGCTGGGGGTTTCGGGCGCACCGGTAGTTCAGG

4301 BGLI, 4308 BALI,

4322 LeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsnCysGly  
CTCACCGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAGAACTGCGGC  
GAGTGGCTCTCCGAAATACAACCCCGGAGAAATGGTTAAGTTCCTCCCTCTTGACGCCG

4345 APAI,

4382 TyrArgArgCysArgAlaSerGlyValLeuThrThrSerCysGlyAsnThrLeuThrCys  
TATCGCAGGTGCCGCGCGAGCGGCGTACTGACAACCTAGCTGTGGTAACACCCTCACTTGC  
ATAGCGTCCACGGCGCGCTCGCCGCATGACTGTTGATCGACACCATTGTGGGAGTGAACG

4442 TyrIleLysAlaArgAlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMetLeuVal  
TACATCAAGGCCCGGGCAGCCTGTGAGCCGCGAGGGCTCCAGGACTGCACCATGCTCGTG  
ATGTAGTTCGGGGCCCGTCCGACAGCTCGGCGTCCCGAGGTCCTGACGTGGTACGAGCAC

4452 SMAI XMAI,

4502 CysGlyAspAspLeuValValIleCysGluSerAlaGlyValGlnGluAspAlaAlaSer  
TGTGGCGACGACTTAGTCGTTATCTGTGAAAGCGCGGGGGTCCAGGAGGACGCGGCGAGC  
ACACCGCTGCTGAATCAGCAATAGACACTTTCGCGCCCCCAGGTCCTCTGCGCCGCTCG

4508 DRD1, 4511 TTH3I,

4562 LeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProGln  
CTGAGAGCCTTCACGGAGGCTATGACCAGGTACTCCGCCCCCCTGGGGACCCCCACAA  
GACTCTCGGAAGTGCTCCGATACTGGTCCATGAGGCGGGGGGACCCCTGGGGGGTGT

4622 ProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaHisAsp  
CCAGAATACGACTTGGAGCTCATAACATCATGCTCCTCCAACGTGTCAGTCGCCCACGAC  
GGTCTTATGCTGAACCTCGAGTATTGTAGTACGAGGAGGTTGCACAGTCAGCGGGTGCTG

4637 SACI,

4682 GlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAlaArgAla  
GGCGCTGGAAAGAGGGTCTACTACCTCACCCGTGACCCTACAACCCCCCTCGCGAGAGCT  
CCGCGACCTTTCTCCAGATGATGGAGTGGGCACTGGGATGTTGGGGGGAGCGCTCTCGA

4731 NRUI,

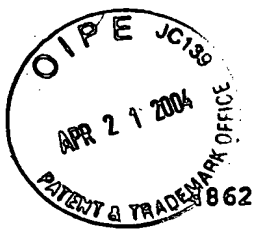
4742 AlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIleMetPhe  
GCGTGGGAGACAGCAAGACACACTCCAGTCAATTCCTGGCTAGGCAACATAATCATGTTT  
CGCACCTCTGTGTTCTGTGTGAGGTCAGTTAAGGACCGATCCGTTGTATTAGTACAA

4802 AlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePheSerValLeuIleAla  
GCCCCACACTGTGGGCGAGGATGATACTGATGACCATTCTTTAGCGTCTTATAGCC  
CGGGGGTGACACCCGCTCCTACTATGACTACTGGGTAAAGAAATCGCAGGAATATCGG

4806 PFLM1, 4807 DRA3,

ArgAspGlnLeuGluGlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSerIleGlu

FIG. 21H



862 AGGGACCAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGGCCTGCTACTCCATAGAA  
TCCCTGGTTCGAACCTTGTCCGGGAGCTAACGCTCTAGATGCCCCGGACGATGAGGTATCTT  
^

4893 BGL2,

4922 ProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSerLeuHis  
CCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTCCTACTCCAC  
GGTGACCTAGATGGAGGTTAGTAAGTTTCTGAGGTACCGGAGTCGCGTAAAGTGAGGTG  
^

4954 NCOI,

4982 SerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGlyValPro  
AGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGCATGCCTCAGAAAACCTGGGGTACCG  
TCAATGAGAGGTCCACTTTAGTTATCCACCGGCGTACGGAGTCTTTTGAACCCCATGGC  
^

5015 SPHI, 5035 KPNI,

5042 ProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAlaArgGly  
CCCTTGCGAGCTTGGAGACACCGGGCCCGAGCGTCCGCGTAGGCTTCTGGCCAGAGGA  
GGGAACGCTCGAACCTCTGTGGCCCGGGCCTCGCAGGCGCGATCCGAAGACCGGTCTCCT  
^

5064 APAI, 5091 BALI,

5102 GlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLys  
GGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAAGCTCAAA  
CCGTCCCGACGGTATACACCGTTCATGGAGAAGTTGACCCGTCATTCTTGTTCGAGTTT  
^

5113 NDEI,

5162 LeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAlaGlyTyr  
CTCACTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCCGGCTGGTTCACGGCTGGCTAC  
GAGTGAGGTTATCGCCGGCGACCGGTTCGACCTGAACAGGCCGACCAAGTGCCGACCGATG  
^ ^ ^

5174 NOTI, 5175 EAG1 XMA3, 5182 BALI, 5186 PVU2,

5222 SerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCys  
AGCGGGGGAGACATTTATCACAGCGTGTCTCATGCCCGGCCCGCTGGATCTGGTTTTGC  
TCGCCCCCTCTGTAAATAGTGTGCGACAGAGTACGGGCCGGGGCGACCTAGACCAAACG  
^

5240 DRA3,

5282 LeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgMetSerThrAsn  
CTACTCCTGCTTGCTGCAGGGGTAGGCATCTACCTCCTCCCAACCGAATGAGCACGAAT  
GATGAGGACGAACGACGTCCCATCCGTAGATGGAGGAGGGGTTGGCTTACTCGTGCTTA  
^

5295 PSTI,

5342 ProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGlnAspValLysPhe  
CCTAAACCTCAAAGAAAGACCAAACGTAACACCAACCGGCGGCCGAGGACGTCAAGTTC  
GGATTTGGAGTTTCTTTCTGGTTTGATTGTGGTTGGCCCGGCGTCTGCAGTTCAAG  
^ ^ ^

5380 NOTI, 5381 EAG1 XMA3, 5390 AAT2, 5401 SMAI XMAI,

5402 ProGlyGlyGlyGlnIleValGlyGlyValTyrLeuLeuProArgArgGlyProArgLeu  
CCGGGTGGCGGTCAGATCGTTGGTGGAGTTTACTTGTGCGCGCAGGGGGCCCTAGATTG  
GGCCACCGCCAGTCTAGCAACCACCTCAAATGAACAACGGCGCGTCCCCGGGATCTAAC  
^

FIG. 21I



5449 APAI,

5462 GlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGlyArgArgGlnPro  
GGTGTGCGCGCGACGAGAAAGACTTCCGAGCGGTCGCAACCTCGAGGTAGACGTCAGCCT  
CCACACGCGCGCTGCTCTTTCTGAAGGCTCGCCAGCGTTGGAGCTCCATCTGCAGTCGGA  
^ ^ ^ ^

5467 BSSH2, 5478 XMNI, 5502 XHOI, 5511 AAT2,

5522 IleProLysAlaArgArgProGluGlyArgThrTrpAlaGlnProGlyTyrProTrpPro  
ATCCCCAAGGCTCGTCGGCCCCGAGGGCAGGACCTGGGCTCAGCCCGGGTACCCTTGGCCC  
TAGGGGTTCCGAGCAGCCGGGCTCCCGTCCTGGACCCGAGTCGGGCCCCATGGGAACCGGG  
^ ^ ^ ^

5548 ALWN1, 5558 ESP1, 5564 SMAI XMAI, 5568 KPNI,

5582 LeuTyrGlyAsnGluGlyCysGlyTrpAlaGlyTrpLeuLeuSerProArgGlySerArg  
CTCTATGGCAATGAGGGCTGCGGGTGGGCGGGATGGCTCCTGTCTCCCCGTGGCTCTCGG  
GAGATACCGTTACTCCCGACGCCCCACCCGCCCTACCGAGGACAGAGGGGCACCGAGAGCC

5642 ProSerTrpGlyProThrAspProArgArgArgSerArgAsnLeuGlyLysValIleAsp  
CCTAGCTGGGGCCCCACAGACCCCGGGCGTAGGTCGCGCAATTTGGGTAAGGTCATCGAT  
GGATCGACCCCGGGGTGTCTGGGGGCCGCATCCAGCGCGTTAAACCCATTCCAGTAGCTA  
^ ^

5650 APAI, 5696 CLAI,

5702 ThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProLeuValOC AM  
ACCCTTACGTGCGGCTTCGCCGACCTCATGGGGTACATACCGCTCGTCTAATAGTCGAC  
TGGGAATGCACGCCGAAGCGGCTGGAGTACCCCATGTATGGCGAGCAGATTATCAGCTG  
^ ^

5724 HGIE2, 5755 SALI,

FIG. 21J



MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsn  
2 AGCTTACAAAACAAAATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTACTCAAC  
TCGAATGTTTTGTTTTACCGACGTATACGTCGAGTCCCGATATTCCACGATCATGAGTTG  
1 HIND3, 24 NDEI, 52 SCAI,  
ProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGlyIleAsp  
62 CCCTCTGTTGCTGCAACACTGGGCTTTGGTGCTTACATGTCCAAGGCTCATGGGATCGAT  
GGGAGACAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCCCTAGCTA  
116 CLAI,  
ProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyrSerThr  
122 CCTAACATCAGGACCGGGGTGAGAACAATTACCACTGGCAGCCCCATCACGTACTCCACC  
GGATTGTAGTCCTGGCCCCACTCTTGTTAATGGTGACCGTCGGGGTAGTGATGAGGTGG  
TyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIleIleCys  
182 TACGGCAAGTTCCTTGCCGACGGCGGGTGCTCGGGGGGCGCTTATGACATAATAATTTGT  
ATGCCGTTCAAGGAACGGCTGCCGCCCCACGAGCCCCCGCGAATACTGTATTATTAAACA  
AspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGln  
242 GACGAGTGCCACTCCACGGATGCCACATCCATCTTGGGCATTGGCACTGTCCTTGACCAA  
CTGCTCACGGTGAGGTGCCTACGGTGTAGGTAGAACCCTAACCCTGACAGGAACCTGGTT  
AlaGluThrAlaGlyAlaArgLeuValValLeuAlaThrAlaThrProProGlySerVal  
302 GCAGAGACTGCGGGGGCGAGACTGGTTGTGCTCGCCACCGCCACCCCTCCGGGGCTCCGTC  
CGTCTCTGACGCCCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCCGAGGCAG  
303 ALWN1,  
ThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIleProPhe  
362 ACTGTGCCCCATCCCAACATCGAGGAGGTTGCTCTGTCCACCACCGGAGAGATCCCTTTT  
TGACACGGGGTAGGGTTGTAGCTCCTCCAACGAGACAGGTGGTGGCCTCTCTAGGGAAAA  
TyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePheCysHis  
422 TACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGAGACATCTCATCTTCTGTCAT  
ATGCCGTTCCGATAGGGGGAGCTTCATTAGTTCCCCCCTCTGTAGAGTAGAAGACAGTA

FIG. 22A



482 SerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsnAlaVal  
TCAAAGAAGAAGTGCGACGAACTCGCCGCAAAGCTGGTCGCATTGGGCATCAATGCCGTG  
AGTTTCTTCTTCACGCTGCTTGAGCGGCGTTTCGACCAGCGTAACCCGTAGTTACGGCAC

542 AlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValValValVal  
GCCTACTACCGCGGTCTTGACGTGTCCGTCATCCCGACCAGCGGCGATGTTGTGTCGTCGTG  
CGGATGATGGCGCCAGAACTGCACAGGCAGTAGGGCTGGTCGCCGCTACAACAGCAGCAC  
550 SAC2, 560 DRD1,

602 AlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsn  
GCAACCGATGCCCTCATGACCGGCTATACCGGCGACTTCGACTCGGTGATAGACTGCAAT  
CGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTGACGTTA  
615 BSPH1,

662 ThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGluThrIle  
ACGTGTGTCAACCCAGACAGTCGATTTTCAGCCTTGACCCTACCTTCACCATTGAGACAATC  
TGCACACAGTGGGTCTGTGAGCTAAAGTCGGAAGTGGGATGGAAGTGGTAACTCTGTTAG

722 ThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLys  
ACGCTCCCCAAGATGCTGTCTCCCGCACTCAACGTCGGGGCAGGACTGGCAGGGGGAAG  
TGCGAGGGGGTTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCCTGACCGTCCCCCTTC

782 ProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSer  
CCAGGCATCTACAGATTTGTGGCACCGGGGAGCGCCCCCTCCGGCATGTTGACTCGTCC  
GGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAGCTGAGCAGG  
816 BGLI, 833 DRD1,

842 ValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThr  
GTCCTCTGTGAGTGCTATGACGCAGGCTGTGCTTGGTATGAGCTCACGCCCGCCGAGACT  
CAGGAGACACTCACGATACTGCGTCCGACACGAACCATACTCGAGTGCGGGCGGCTCTGA  
881 SACI,

902 ThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeu  
ACAGTTAGGCTACGAGCGTACATGAACACCCCGGGGCTTCCCGTGTGCCAGGACCATCTT  
TGTCATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCACACGGTCTTGCTAGAA  
931 SMAI XMAI,

962 GluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGln  
GAATTTTGGGAGGGCGTCTTTACAGGCCTCACTCATATAGATGCCACTTTCTATCCCAG  
CTTAAAACCCCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGATAGGGTC  
985 STUI,

1022 ThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAla  
ACAAAGCAGAGTGGGGAGAACCTTCCTTACCTGGTAGCGTACCAAGCCACCGTGTGCGCT  
TGTTTCGTCTCACCCCTCTTGGAAGGAATGGACCATCGCATGGTTCGGTGGCACACGCGA  
1069 DRA3,

1082 ArgAlaGlnAlaProProProS rTrpAspGlnMetTrpLysCysLeuIleArgLeuLys  
AGGGCTCAAGCCCCCTCCCCATCGTGGGACCAGATGTGGAAGTGTGTTGATTGCGCTCAAG

FIG. 22B



TCCCGAGTTCGGGGAGGGGGTAGCACCCCTGGTCTACACCTTCACAACTAAGCGGAGTTG

1142 Pr ThrLeuHisGlyProThrPr LeuLeuTyrArgLeuGlyAlaValGlnAsnGluIle  
CCCACCCCTCCATGGGCCAACACCCCTGCTATACAGACTGGGCGCTGTTTCTCAGAAATGAAATC  
GGGTGGGAGGTACCCGGTTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTACTTTAG  
^  
1150 NCOI,

1202 ThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeuGluVal  
ACCCTGACGCACCCAGTCACCAAATACATCATGACATGCATGTCGGCCGACCTGGAGGTC  
TGGGACTGCGTGGGTGAGTGGTTTATGTAGTACTGTACGTACAGCCGGCTGGACCTCCAG  
^ ^ ^ ^ ^  
1230 BSPH1, 1234 DRD1, 1237 AVA3, 1245 EAG1 XMA3, 1250 DRD1,

1262 ValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyrCysLeu  
GTCACGAGCACCTGGGTGCTCGTTGGCGGCGTCTGGCTGCTTTGGCCGCGTATTGCCTG  
CAGTGCTCGTGGACCCACGAGCAACCGCCGAGGACCGACGAAACCGGCGCATAACGGAC

1322 SerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAlaIleIle  
TCAACAGGCTGCGTGGTCATAGTGGGCAGGGTCTGCTTGTCCGGGAAGCCGGCAATCATA  
AGTTGTCCGACGCACCGAGTATCACCCGTCCAGCAGAACAGGCCCTTCGGCCGTTAGTAT  
^  
1369 NAEI,

1382 ProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGlnHisLeu  
CCTGACAGGGAAGTCTCTACCGAGAGTTTCATGAGATGGAAGAGTGCTCTCAGCACTTA  
GGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCAGGAGAGTCGTGAAT  
^  
1385 DRD1,

1442 ProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeuGlyLeu  
CCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGCAGAAGGCCCTCGGCCTC  
GGCATGTAGCTCGTTCCTACTACGAGCGGCTCGTCAAGTTCGTCTTCCGGGAGCCGGAG

1502 LeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsnTrpGln  
CTGCAGACCGCGTCCCGTCAGGCAGAGGTTATCGCCCCCTGCTGTCCAGACCAACTGGCAA  
GACGTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTGACCGTT  
^ ^ ^  
1502 PSTI, 1507 TTH3I,

1562 LysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeu  
AAACTCGAGACCTTCTGGGCGAAGCATATGTGGAACCTCATCAGTGGGATACAATACTTG  
TTTGAGCTCTGGAAGACCCGCTTCGTATACACCTTGAAGTAGTCACCCTATGTTATGAAC  
^ ^  
1565 XHOI, 1586 NDEI,

1622 AlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPheThrAla  
GCGGGCTTGTCAACGCTGCCTGGTAACCCCGCCATTGCTTCATTGATGGCTTTTACAGCT  
CGCCCGAACAGTTGCGACGGACCATTTGGGGCGGTAACGAAGTAACCTACCGAAAATGTCCA  
^ ^  
1643 BSTE2, 1677 ALWN1 PVU2,

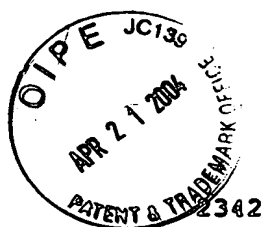
1682 AlaValThrSerPr LeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGlyGlyTrp  
GCTGTCAACAGCCCACTAACCCTAGCCAAACCCCTCTTCAACATATTGGGGGGGTGG  
CGACAGTGGTGGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCCCCACC

FIG. 22C



1742 ValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGlyL uAla  
GTGGCTGCCCAGCTCGCCGCCCCCGGTGCCGCTACTGCCTTTGTGGGCGCTGGCTTAGCT  
CACCGACGGGTGAGCGGCGGGGGCCACGGCGATGACGGAAACACCCGCGACCGAATCGA  
1794 ESP1,  
1802 GlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyr  
GGCGCCGCCATCGGCAGTGTGGACTGGGGAAGGTCTCATAGACATCCTTGCAGGGTAT  
CCGCGGCGGTAGCCGTCACAACCTGACCCCTTCCAGGAGTATCTGTAGGAACGTCCATA  
1802 KAS1 NARI,  
1862 GlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluValProSer  
GGCGCGGGCGTGGCGGGAGCTCTTGTGGCATTCAAGATCATGAGCGGTGAGGTCCCCTCC  
CCGCGCCCGCACCGCCCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAGGGGAGG  
1878 SACI, 1899 BSPH1,  
1922 ThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuValValGly  
ACGGAGGACCTGGTCAATCTACTGCCCCGCCATCCTCTCGCCCGGAGCCCTCGTAGTCGGC  
TGCCTCCTGGACCAGTTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCATCAGCCG  
1928 TTH3I,  
1982 ValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrp  
GTGGTCTGTGCAGCAATACTGCGCCGGCACGTTGGCCCCGGGCGAGGGGGCAGTGCAGTGG  
CACCAGACACGTCGTTATGACCGGCCGTGCAACCGGGCCCGCTCCCCCGTCACGTCACC  
2004 NAEI, 2017 SMAI XMAI,  
2042 MetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrVal  
ATGAACCGGCTGATAGCCTTCGCCTCCCGGGGGAACCATGTTCCCCCACGCACTACGTG  
TACTTGGCCGACTATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGGTGCGTGATGCAC  
2067 SMAI XMAI, 2093 DRA3,  
2102 ProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGln  
CCGGAGAGCGATGCAGCTGCCCGCGTCACTGCCATACTCAGCAGCCTCACTGTAACCCAG  
GGCCTCTCGCTACGTCGACGGGCGCAGTGACGGTATGAGTCGTCGGAGTGACATTGGGTC  
2115 PVU2, 2159 ALWN1,  
2162 LeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSerGlySer  
CTCCTGAGGCGACTGCACCACTGGATAAGCTCGGAGTGTAACCTCCATGCTCCGGTTCC  
GAGGACTCCGCTGACGTGGTCACCTATTCGAGCCTCACATGGTGAGGTACGAGGCCAAGG  
2164 MST2, 2220 ECON1,  
2222 TrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeu  
TGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCGACTTTAAGACCTGGCTA  
ACCGATTCCCTGTAGACCCTGACCTATACGCTCCACAACCTCGCTGAAATTCTGGACCGAT  
2282 LysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyr  
AAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCCTTTGTGTCTGCCAGCGCGGGTAT  
TTTCGATTTCGAGTACGGTGTGACGCGACCCTAGGGGAAACACAGGACGGTCGCGCCATA  
2285 ESP1, 2300 PVU2, 2310 BAMHI,

FIG. 22D



2342 LysGlyValTrpArgGlyAspGlyIleMetHisThrArgCysHisCysGlyAlaGluIle  
AAGGGGGTCTGGCGAGGGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCTGAGATC  
TTCCCCCAGACCGCTCCCCTGCCGTAGTACGTGTGAGCGACGGTGACACCTCGACTCTAG

2402 ThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArgAsnMet  
ACTGGACATGTCAAAAACGGGACGATGAGGATCGTCGGTCCTAGGACCTGCAGGAACATG  
TGACCTGTACAGTTTTTGGCCCTGCTACTCCTAGCAGCCAGGATCCTGGACGTCCTTGATC  
^ ^ ^

2425 BSAB1, 2441 AVR2, 2448 SSE83871, 2449 PSTI,

2462 TrpSerGlyThrPheProIleAsnAlaTyrThrThrGlyProCysThrProLeuProAla  
TGGAGTGGGACCTTCCCCATTAATGCCTACACCACGGGCCCCCTGTACCCCCCTTCCTGCG  
ACCTCACCTGGAAGGGGTAATTACGGATGTGGTGCCCGGGACATGGGGGGAAGGACGC  
^ ^

2480 ASE1, 2497 APAI,

2522 ProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIleArgGln  
CCGAACTACACGTTTCGCGCTATGGAGGGTGTCTGCAGAGGAATACGTGGAGATAAGGCAG  
GGCTTGATGTGCAAGCGCGATACCTCCCACAGACGTCTCCTTATGCACCTCTATTCCGTC  
^

2553 PSTI,

2582 ValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysProCysGln  
GTGGGGGACTTCCACTACGTGACGGGTATGACTACTGACAATCTTAAATGCCCGTGCCAG  
CACCCCTGAAGGTGATGCACTGCCCATCTGATGACTGTTAGAATTTACGGGCACGGTC  
^

2594 DRA3,

2642 ValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPheAlaPro  
GTCCCATCGCCCGAATTTTTCACAGAATTGGACGGGGTGCGCCTACATAGGTTTGCGCCC  
CAGGGTAGCGGGCTTAAAAAGTGTCTTAACCTGCCCCACGCGGATGTATCCAAACGCGGG

2702 ProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGluTyrPro  
CCCTGCAAGCCCTTGCTGCGGGAGGAGGTATCATTAGAGTAGGACTCCACGAATACCCG  
GGGACGTTGCGGAACGACGCCCTCCTCCATAGTAAGTCTCATCCTGAGGTGCTTATGGGC  
^

2757 HGIE2,

2762 ValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSerMetLeu  
GTAGGGTCGCAATTACCTTGCGAGCCCGAACCAGGACGTGGCCGTGTTGACGTCCATGCTC  
CATCCCAGCGTTAATGGAACGCTCGGGCTTGGCCTGCACCGGCACAACCTGCAGGTACGAG  
^

2809 AAT2,

2822 ThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGlySerPro  
ACTGATCCCTCCCATATAACAGCAGAGGCGGCGGGCGAAGGTTGGCGAGGGGATACCC  
TGACTAGGGAGGGTATATTGTCGTCTCCGCGGCCCCGCTCCAACCGCTCCCTAGTGGG  
^

2850 EAG1 XMA3,

2882 ProSerValAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuLysAlaThrCys  
CCCTCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCTCAAGGCAACTTGC  
GGGAGACACCGGTGAGGAGCCGATCGGTGATAGGCGAGGTAGAGAGTTCCGTTGAACG  
^ ^

2889 BALI, 2903 NHEI,

FIG. 22E



2942 ThrAlaAsnHisAspS rProAspAlaGluLeuIleGluAlaAsnLeuLeuTrpArgGln  
ACCGCTAACCATGACTCCCTGATGCTGAGCTCATAGAGGCCAACCTCCTATGGAGGCAG  
TGGCGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGGTTGGAGGATACCTCCGTC  
2966 ESP1, 2969 SACI,  
3002 GluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeuAspSer  
GAGATGGGCGGCAACATCACCAGGGTTGAGTCAGAAAACAAAGTGGTGATTCTGGACTCC  
CTCTACCCGCCGTTGTAGTGGTCCCAACTCAGTCTTTTGTTCACCACTAAGACCTGAGG  
3062 PheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGluIleLeu  
TTCGATCCGCTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGCAGAAATCCTG  
AAGCTAGGCGAACACCGCCTCCTCCTGCTCGCCCTCTAGAGGCATGGGCGTCTTTAGGAC  
3096 BGL2,  
3122 ArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyrAsnPro  
CGGAAGTCTCGGAGATTGCGCCAGGCCCTGCCCGTTTGGGCGCGGCCGACTATAACCCC  
GCCTTCAGAGCCTCTAAGCGGGTCCGGGACGGGCAAACCCGCGCCGCGCTGATATTGGGG  
3143 ALWN1, 3164 EAG1 XMA3,  
3182 ProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGlyCysPro  
CCGCTAGTGGAGACGTGGAAAAAGCCCGACTACGAACCACCTGTGGTCCATGGCTGCCCCG  
GGCGATCACCTCTGCACCTTTTTTCGGGCTGATGCTTGGTGGACACCAGGTACCGACGGGG  
3217 HGIE2, 3229 NCOI,  
3242 LeuProProProLysSerProProValProProProArgLysLysArgThrValValLeu  
CTTCCACCTCCAAAGTCCCTCCTGTGCCTCCGCCTCGGAAGAAGCGGACGGTGGTCCTC  
GAAGGTGGAGGTTTCAGGGGAGGACACGGAGGCGGAGCCTTCTTCGCCTGCCACCAGGAG  
3302 ThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGlySerSer  
ACTGAATCAACCCTATCTACTGCCTTGGCCGAGCTCGCCACCAGAAGCTTTGGCAGCTCC  
TGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGGTCTTCGAAACCGTCGAGG  
3332 SACI, 3346 HIND3,  
3362 SerThrSerGlyIleThrGlyAspAsnThrThrThrSerSerGluProAlaProSerGly  
TCAACTTCCGGCATTACGGGCGACAATACGACAACATCCTCTGAGCCCCGCCCTTCTGGC  
AGTTGAAGGCCGTAATGCCCGCTGTTATGCTGTTGTAGGAGACTCGGGCGGGGAAGACCG  
3422 CysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGlyGluPro  
TGCCCCCCCCGACTCCGACGCTGAGTCCTATTCTCCATGCCCCCCTGGAGGGGGAGCCT  
ACGGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGGACCTCCCCCTCGGA  
3437 EAM11051,  
3482 GlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsnAlaGlu  
GGGGATCCGGATCTTAGCGACGGGTCAACGGTCAGTAGTGAGGCCAACGCGGAG  
CCCCTAGGCCTAGAATCGCTGCCCAGTACCAGTTGCCAGTCATCACTCCGGTTGCGCCTC  
3484 BAMHI, 3485 BSAB1, 3487 BSPE1,  
3542 AspValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuValThrProCysAla  
GATGTCGTGTGCTCAATGTCTTACTCTTGGACAGGCGCACTCGTCACCCCGTGCGCC  
CTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCGCGTGAGCAGTGGGGCACGCGG

FIG. 22F



3589 DRA3, 3600 SAC2,

3602 AlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsn  
GCGGAAGAACAGAACTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTCACCACAAT  
CGCCTTCTTGCTTTGACGGGTAGTTACGTGATTGTTGAGCAACGATGCAGTGGTGTTA  
^ ^

3611 ALWN1, 3655 PFLM1,

3662 LeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAsp  
TTGGTGTATTCCACCACCTCACGCAGTGCTTGCCAAAGGCAGAAGAAAGTCACATTTGAC  
AACCACATAAGGTGGTGGAGTGCCTCACGAACGGTTTCCGTCTTCTTTCAGTGTAAGTGC  
^

3681 DRA3,

3722 ArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAlaAlaAla  
AGACTGCAAGTTCTGGACAGCCATTACCAGGACGTACTCAAGGAGGTTAAAGCAGCGGCG  
TCTGACGTTCAAGACCTGTCGGTAATGGTCTGTCATGAGTTCCTCCAATTTTCGTCGCCGCG

3782 SerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrProProHis  
TCAAAGTGAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGCAGCCTGACGCCCCCACAC  
AGTTTTCACTTCCGATTGAACGATAGGCATCTCCTTGAACGTCGGACTGCGGGGGTGTG  
^

3816 HIND3,

3842 SerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAla  
TCAGCCAAATCCAAGTTTGGTTATGGGGCAAAAGACGTCCGTTGCCATGCCAGAAAGGCC  
AGTCGGTTTAGGTTCAAACCAATACCCCGTTTTCTGCAGGCAACGGTACGGTCTTTCCGG  
^ ^

3875 AAT2, 3890 BGLI,

3902 ValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrProIleAsp  
GTAACCCACATCAACTCCGTGTGGAAGACCTTCTGGAAGACAATGTAACACCAATAGAC  
CATTGGGTGTAGTTGAGGCACACCTTCTGGAAGACCTTCTGTTACATTGTGGTTATCTG

3962 ThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGlyArgLys  
ACTACCATCATGGCTAAGAACGAGGTTTTCTGCGTTCAGCCTGAGAAGGGGGGTCGTAAG  
TGATGGTAGTACCGATTCTTGCTCCAAAGACGCAAGTCGGACTCTTCCCCCAGCATTC

4022 ProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMetAlaLeu  
CCAGCTCGTCTCATCGTGTTCCTCCGATCTGGGCGTGCGCGTGTGCGAAAAGATGGCTTTG  
GGTCGAGCAGAGTAGCACAAGGGGCTAGACCCGCACGCGCACACGCTTTTCTACCGAAAC

4082 TyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPheGlnTyr  
TACGACGTGGTTACAAAGCTCCCCTTGGCCGTGATGGGAAGCTCCTACGGATTCCAATAC  
ATGCTGCACCAATGTTTTGAGGGGAACCGGCACTACCCTTCGAGGATGCCTAAGGTTATG

4142 SerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThrProMet  
TCACCAGGACAGCGGGTTGAATTCCTCGTGCAAGCGTGGAAGTCCAAGAAAACCCCAATG  
AGTGGTCTGTGCGCCAACTTAAGGAGCACGTTGCGACCTTCAGGTTCTTTTGGGGTTAC  
^

4160 ECORI,

4202 GlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIleArgThr  
GGGTTCTCGTATGATACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATCCGTACG  
CCCAAGAGCATACTATGGGCGACGAACTGAGGTGTCACTGACTCTCGCTGTAGGCATGC  
^ ^

FIG. 22G



4229 DRD1, 4236 ALWN1,

4262 GluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIleLysSer  
GAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCAAGCCCGCTGGCCATCAAGTCC  
CTCCTCCGTTAGATGGTTACAACACTGGAGCTGGGGGTTCTGGGCGCACCGGTAGTTCAGG  
^ ^

4301 BGLI, 4308 BALI,

4322 LeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsnCysGly  
CTCACCAGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAGAACTGCGGC  
GAGTGGCTCTCCGAAATACAACCCCGGGAGAATGGTTAAGTTCCTCCCTCTTGACGCCG  
^

4345 APAI,

4382 TyrArgArgCysArgAlaSerGlyValLeuThrThrSerCysGlyAsnThrLeuThrCys  
TATCGCAGGTGCCGCGCGAGCGCGTACTGACAACCTAGCTGTGGTAACACCCTCACTTGC  
ATAGCGTCCACGGCGCGCTCGCCGCATGACTGTTGATCGACACCATTTGTGGGAGTGAACG

4442 TyrIleLysAlaArgAlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMetLeuVal  
TACATCAAGGCCCGGGCAGCCTGTGAGCCGCGAGGGCTCCAGGACTGCACCATGCTCGTG  
ATGTAGTTCCGGGCCCGCTCGGACAGCTCGGCGTCCCGAGGTCTGACGTGGTACGAGCAC  
^

4452 SMAI XMAI,

4502 CysGlyAspAspLeuValValIleCysGluSerAlaGlyValGlnGluAspAlaAlaSer  
TGTGGCGACGACTTAGTCGTTATCTGTGAAAGCGCGGGGGTCCAGGAGGACGCGCGCAGC  
ACACCGCTGCTGAATCAGCAATAGACACTTTCGCGCCCCCAGGTCTCTGCGCCGCTCG  
^ ^

4508 DRD1, 4511 TTH3I,

4562 LeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProGln  
CTGAGAGCCTTCACGGAGGCTATGACCAGGTACTCCGCCCCCTGGGGACCCCCACAA  
GACTCTCGGAAGTGCTCCGATACTGGTCCATGAGGCGGGGGGACCCCTGGGGGGTGT

4622 ProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaHisAsp  
CCAGAATACGACTTGGAGCTCATAACATCATGCTCCTCCAACGTGTCAGTCGCCCACGAC  
GGTCTTATGCTGAACCTCGAGTATTGTAGTACGAGGAGGTTGCACAGTCAGCGGGTGCTG  
^

4637 SACI,

4682 GlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAlaArgAla  
GGCGCTGGAAAGAGGGTCTACTACCTACCCGTGACCCTACAACCCCTCGCGAGAGCT  
CCGCGACCTTTCTCCAGATGATGGAGTGGGCACTGGGATGTTGGGGGGAGCGCTCTCGA  
^

4731 NRUI,

4742 AlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIleMetPhe  
GCGTGGGAGACAGCAAGACACTCCAGTCAATTCCTGGCTAGGCAACATAATCATGTTT  
CGCACCTCTGTGTTCTGTGTGAGGTCAGTTAAGGACCGATCCGTTGTATTAGTACAAA

4802 AlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePheSerValLeuIleAla  
GCCCCACACTGTGGGCGAGGATGATACTGATGACCCATTTCTTTAGCGTCCTTATAGCC  
CGGGGGTGTGACACCCGCTCCTACTATGACTACTGGGTAAAGAAATCGCAGGAATATCGG  
^ ^

4806 PFLM1, 4807 DRA3,

ArgAspGlnLeuGluGlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSerIleGlu

FIG. 22H



4862 AGGGACCAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGGCCTGCTACTCCATAGAA  
TCCCTGGTCTGAACCTTGTCCGGGAGCTAACGCTCTAGATGCCCCGGACGATGAGGTATCTT  
^

4893 BGL2,

4922 ProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSerLeuHis  
CCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTTCACTCCAC  
GGTGACCTAGATGGAGGTTAGTAAGTTTCTGAGGTACCGGAGTCGCGTAAAAGTGAGGTG  
^

4954 NCOI,

4982 SerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGlyValPro  
AGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGCATGCCTCAGAAAACCTTGGGGTACCG  
TCAATGAGAGGTCCACTTTAGTTATCCCACGGCGTACGGAGTCTTTTGAACCCCATGGC  
^

5015 SPHI, 5035 KPNI,

5042 ProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAlaArgGly  
CCCTTGCGAGCTTGAGACACCGGGCCCGAGCGTCCGCGCTAGGCTTCTGGCCAGAGGA  
GGGAACGCTCGAACCTCTGTGGCCCCGGGCTCGCAGGCGCGATCCGAAGACCGGTCTCCT  
^

5064 APAI, 5091 BALI,

5102 GlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLys  
GGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAAGCTCAAA  
CCGTCCCAGCGTATACACCGTTCATGGAGAAGTTGACCCGTCATTCTTGTTTCGAGTTT  
^

5113 NDEI,

5162 LeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAlaGlyTyr  
CTCACTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCCGGCTGGTTCACGGCTGGCTAC  
GAGTGAGGTTATCGCCGGCGACCGGTGACCTGAACAGGCCGACCAAGTGCCGACCGATG  
^ ^ ^

5174 NOTI, 5175 EAG1 XMA3, 5182 BALI, 5186 PVU2,

5222 SerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCys  
AGCGGGGGAGACATTTATCACAGCGTGTCTCATGCCCGGCCCCGCTGGATCTGGTTTTGC  
TCGCCCCCTCTGTAAATAGTGTCGCACAGAGTACGGGCGGGGCGACCTAGACCAAACG  
^

5240 DRA3,

5282 LeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgMetSerThrAsn  
CTACTCCTGCTTGCTGCAGGGGTAGGCATCTACCTCCTCCCCAACCGAATGAGCACGAAT  
GATGAGGACGAACGACGTCCCCATCCGTAGATGGAGGAGGGGTGGCTTACTCGTGCTTA  
^

5295 PSTI,

5342 ProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGlnAspValLysPhe  
CCTAAACCTCAAAGAAAGACCAAACGTAACACCAACCGGCGCGCAGGACGTCAAGTTC  
GGATTTGGAGTTTCTTTCTGGTTTGCATTGTGGTTGGCCGCGCGCTCCTGCAGTTCAAG  
^ ^ ^

5380 NOTI, 5381 EAG1 XMA3, 5390 AAT2, 5401 SMAI XMAI,

5402 ProGlyGlyGlyGlnIleValGlyGlyValTyrLeuLeuProArgArgGlyProArgLeu  
CCGGGTGGCGGTGAGATCGTTGGTGGAGTTTACTTGTGCGCGCAGGGGCCCTAGATTG  
GGCCACCGCCAGTCTAGCAACCACCTCAAATGAACAACGGCGCGTCCCCGGGATCTAAC  
^



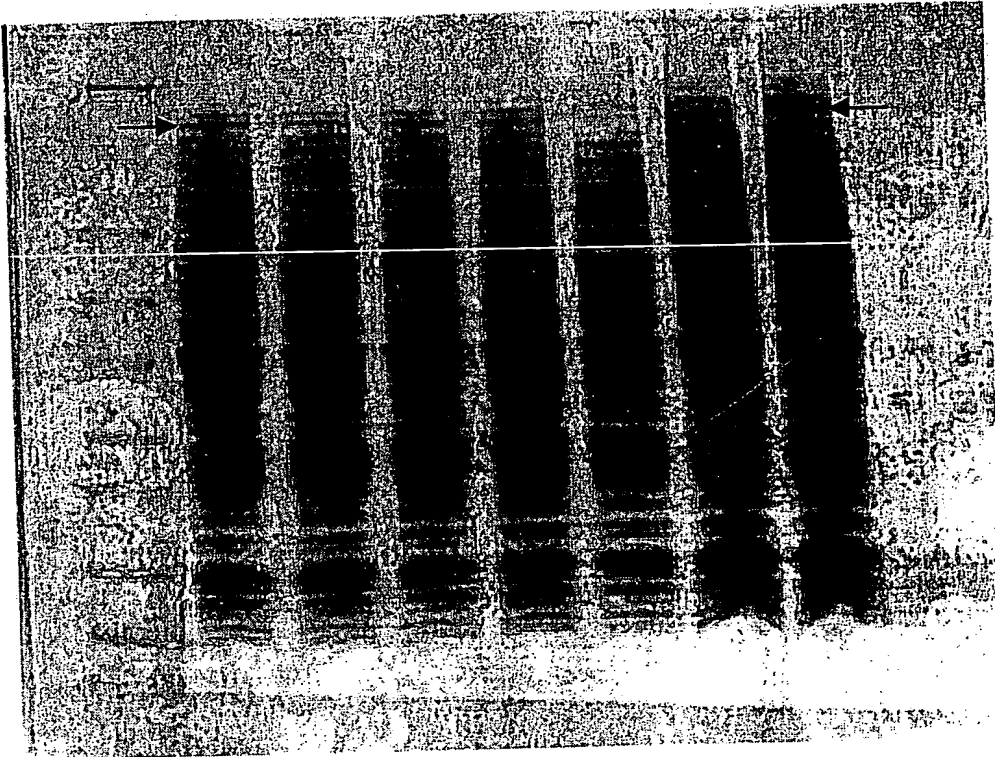


FIG. 23